The Complexity of Approximating Vertex Expansion

Anand Louis [*]
Georgia Tech
anandl@gatech.edu

Prasad Raghavendra [†]
UC Berkeley
prasad@cs.berkeley.edu

Santosh Vempala [*]
Georgia Tech
vempala@cc.gatech.edu

Abstract

We study the complexity of approximating the vertex expansion of graphs $G = (V, E)$, defined as

$$\phi^{\mathsf{V}} \stackrel{\text{def}}{=} \min_{S \subset V} n \cdot \frac{|N(S)|}{|S||V \setminus S|}.$$

We give a simple polynomial-time algorithm for finding a subset with vertex expansion $O(\sqrt{\text{OPT} \log d})$ where d is the maximum degree of the graph. Our main result is an asymptotically matching lower bound: under the Small Set Expansion (SSE) hypothesis, it is hard to find a subset with expansion less than $C\sqrt{\text{OPT} \log d}$ for an absolute constant C. In particular, this implies for all constant $\varepsilon > 0$, it is SSE-hard to distinguish whether the vertex expansion $< \varepsilon$ or at least an absolute constant. The analogous threshold for edge expansion is $\sqrt{\text{OPT}}$ with no dependence on the degree; thus our results suggest that vertex expansion is harder to approximate than edge expansion. In particular, while Cheeger's algorithm can certify constant edge expansion, it is SSE -hard to certify constant vertex expansion in graphs.

Our proof is via a reduction from the *Unique Games* instance obtained from the SSE hypothesis to the vertex expansion problem. It involves the definition of a smoother intermediate problem we call Balanced Analytic Vertex Expansion which is representative of both the vertex expansion and the conductance of the graph. Both reductions (from the UGC instance to this problem and from this problem to vertex expansion) use novel proof ideas.

1   Introduction

Vertex expansion is an important parameter associated with a graph, one that has played a major role in both algorithms and complexity. Given a graph $G = (V, E)$, the vertex expansion of a set $S \subseteq V$ of vertices is defined as

$$\phi^{\mathsf{V}}(S) \stackrel{\text{def}}{=} |V| \cdot \frac{|N(S)|}{|S||V \setminus S|}$$

Here $N(S)$ denotes the outer boundary of the set S, i.e. $N(S) = \{i \in V \setminus S | \exists u \in S \text{ such that } \{u, v\} \in E\}$. The vertex expansion of the graph is given by $\phi^{\mathsf{V}} \stackrel{\text{def}}{=} \min_{S \subset V} \phi^{\mathsf{V}}(S)$. The problem of computing ϕ^{V} is a major primitive for many graph algorithms specifically for those that are based on the divide and conquer paradigm [LR99]. It is NP-hard to compute the vertex expansion ϕ^{V} of a graph exactly. In this work, we study the approximability of vertex expansion ϕ^{V} of a graph.

A closely related notion to vertex expansion is that of edge expansion. The edge expansion of a set S is defined as

$$\phi(S) \stackrel{\text{def}}{=} \frac{\mu(E(S, \bar{S}))}{\mu(S)}$$

and the edge expansion of the graph is $\phi = \min_{S \subset V} \phi(S)$. Graph expansion problems have received much attention over the past decades, with applications to many algorithmic problems, to the construction of pseudorandom objects and more recenlty due to their connection to the unique games conjecture.

The problem of approximating edge or vertex expansion can be studied at various regimes of parameters of interest. Perhaps the simplest possible version of the problem is to distinguish whether a given graph is an expander. Fix an absolute constant δ_0. A graph is a δ_0-vertex (edge) expander if its vertex (edge) expansion is at least δ_0. The problem of recognizing a vertex expander can be stated as follows:

Definition 1.1. Given a graph G, distinguish between the following two cases

- (Non-Expander) the vertex expansion is $< \varepsilon$

- (Expander) the vertex expansion is $> \delta_0$ for some absolute constant δ_0.

Similarly, one can define the problem of recognizing an edge expander graph.

Notice that if there is some sufficiently small absolute constant ε (depending on δ_0), for which the above problem is easy, then we could argue that it is easy to "recognize" a vertex expander. For the edge case, the Cheeger's inequality yields an algorithm to recognize an edge expander. In fact, it is possible to distinguish a δ_0 edge expander graph, from a graph whose edge expansion is $< \delta_0^2/2$, by just computing the second eigenvalue of the graph Laplacian.

It is natural to ask if there is an efficient algorithm with an analogous guarantee for vertex expansion. More precisely, is there some sufficiently small ε (an arbitrary function of δ_0), so that one can efficiently distinguish between a graph with vertex expansion $> \delta_0$ from one with vertex expansion $< \varepsilon$. In this work, we show a hardness result suggesting that there is no efficient algorithm to recognize vertex expanders. More precisely, our main result is a hardness for the problem of approximating vertex expansion in graphs of bounded degree d. The hardness result shows that the approximability of vertex expansion degrades with the degree, and therefore the problem of recognizing expanders is hard for sufficiently large degree. Furthermore, we exhibit an approximation algorithm for vertex expansion whose guarantee matches the hardness result up to constant factors.

Related Work.   The first approximation for conductance was obtained by discrete analogues of the Cheeger inequality shown by Alon-Milman [AM85] and Alon [Alo86]. Specifically, Cheeger's inequality relates the conductance Φ to the second eigenvalue of the adjacency matrix of the graph – an efficiently computable quantity. This yields an approximation algorithm for Φ, one that is used heavily in practice for graph partitioning. However, the approximation for Φ obtained via Cheeger's inequality is poor in terms of a approximation ratio, especially when the value of Φ is small. An $O(\log n)$ approximation algorithm for Φ was obtained by Leighton and Rao [LR99]. Later work by Linial et al. [LLR95] and Aumann and Rabani [AR98] established a strong connection between the SPARSEST CUT problem and the theory of metric spaces, in turn spurring a large and rich body of literature. The current best algorithm for the problem is an $O(\sqrt{\log n})$ approximation for due to Arora et al. [ARV04] using semidefinite programming techniques.

Ambühl, Mastrolilli and Svensson [AMS07] showed that ϕ^V and ϕ have no PTAS assuming that SAT does not have sub-exponential time algorithms. The current best approximation factor for ϕ^V is $(\sqrt{\log n} \log \log n)$ obtained using a convex relaxation [FHL08]. Beyond this, the situation is much less clear for the approximability of vertex expansion. Applying Cheeger's method leads to a bound of $O(\sqrt{d\,\mathsf{OPT}})$ [Alo86] where d is the maximum degree of the input graph.

Small Set Expansion Hypothesis.   A more refined measure of the edge expansion of a graph is its expansion profile. Specifically, for a graph G the expansion profile is given by the curve

$$\phi(\delta) = \min_{\mu(S) \leqslant \delta} \phi(S) \qquad \forall \delta \in [0, 1/2] \,.$$

The problem of approximating the expansion profile has received much less attention, and is seemingly far less tractable. In summary, the current state-of-the-art algorithms for approximating the expansion profile of a graph are still far from satisfactory. Specifically, the following hypothesis is consistent with the known algorithms for approximating expansion profile.

Hypothesis (Small-Set Expansion Hypothesis, [RS10])**.** *For every constant $\eta > 0$, there exists sufficiently small $\delta > 0$ such that given a graph G it is NP-hard to distinguish the cases,*

YES: *there exists a vertex set S with volume $\mu(S) = \delta$ and expansion $\phi(S) \leqslant \eta$,*

No: *all vertex sets S with volume $\mu(S) = \delta$ have expansion $\phi(S) \geqslant 1 - \eta$.*

Apart from being a natural optimization problem, the SMALL-SET EXPANSION problem is closely tied to the Unique Games Conjecture. Recent work by Raghavendra-Steurer [RS10] established reduction from the SMALL-SET EXPANSION problem to the well known Unique Games problem, thereby showing that Small-Set Expansion Hypothesis implies the Unique Games Conjecture. This result suggests that the problem of approximating expansion of small sets lies at the combinatorial heart of the Unique Games problem.

In a breakthrough work, Arora, Barak, and Steurer [ABS10] showed that the problem SMALL-SET EXPANSION(η, δ) admits a subexponential algorithm, namely an algorithm that runs in time $\exp(n^\eta/\delta)$. However, such an algorithm does not refute the hypothesis that the problem SMALL-SET EXPANSION(η, δ) might be hard for every constant $\eta > 0$ and sufficiently small $\delta > 0$.

The Unique Games Conjecture is not known to imply hardness results for problems closely tied to graph expansion such as BALANCED SEPARATOR. The reason being that the hard instances of these problems are required to have certain global structure namely expansion. Gadget reductions from a unique games instance preserve the global properties of the unique games instance such as lack of expansion. Therefore, showing hardness for graph expansion problems often required a stronger version of the EXPANDING UNIQUE GAMES,

where the instance is guaranteed to have good expansion. To this end, several such variants of the conjecture for expanding graphs have been defined in literature, some of which turned out to be false [AKK⁺08]. The Small-Set Expansion Hypothesis could possibly serve as a natural unified assumption that yields all the implications of expanding unique games and, in addition, also hardness results for other fundamental problems such as BALANCED SEPARATOR. In fact, Raghavendra, Steurer and Tulsiani [RST12] show that the the SSE hypothesis implies that the Cheeger's algorithm yields the best approximation for the balanced separator problem.

Formal Statement of Results. Our first result is a simple polynomial-time algorithm to obtain a subset of vertices S whose vertex expansion is at most $O(\sqrt{\text{OPT} \log d})$. Here d is the largest vertex degree of G. The algorithm is based on a Poincairé-type graph parameter called λ_∞ defined by Bobkov, Houdré and Tetali [BHT00], which approximates ϕ^V. While λ_∞ also appears to be hard to compute, its natural SDP relaxation gives a bound that is within $O(\log d)$, as observed by Steurer and Tetali [ST12], which inspires our first Theorem.

Theorem 1.2. *There exists a polynomial time algorithm which given a graph $G = (V, E)$ having vertex degrees at most d, outputs a set $S \subset V$, such that $\phi^V(S) = O(\sqrt{\text{OPT} \log d})$.*

It is natural to ask if one can prove better inapproximability results for vertex expansion than those that follow from the inapproximability results for edge expansion. Indeed, the best one could hope for would be a lower bound matching the upper bound in the above theorem. Our main result is a reduction from SSE to the problem of distinguishing between the case when vertex expansion of the graph is at most ε and the case when the vertex expansion is at least $\Omega(\sqrt{\varepsilon \log d})$. This immediately implies that it is SSE -hard to find a subset of vertex expansion less than $C\sqrt{\text{OPT} \log d}$ for some constant C. To the best of our knowledge, our work is the first evidence that vertex expansion might be harder to approximate than edge expansion. More formally, we state our main theorem below.

Theorem 1.3. *For every $\varepsilon, \eta > 0$, there exists an absolute constant C such that it is SSE -hard to distinguish between the following two cases for a given graph G with maximum degree d.*

YES : *There exists a set $S \subset V$ of size $|S| = |V|/2$ such that*

$$\phi^V(S) \leqslant \varepsilon$$

No : *For all sets $S \subset V$,*

$$\phi^V(S) \geqslant \min(10^{-10}, C\sqrt{\varepsilon \log d}) - \eta$$

By a suitable choice of parameters in the above theorem, we obtain the main theorem of this work, Theorem 1.3.

Theorem 1.4. *There exists an absolute constant $\delta_0 > 0$ such that for every constant $\varepsilon > 0$ the following holds: Given a graph G, it is SSE -hard to distinguish between the following two cases:*

YES : *There exists a set $S \subset V$ of size $|S| = |V|/2$ such that $\phi^V(S) \leqslant \varepsilon$*

No : *(G is a vertex expander with constant expansion) For all sets $S \subset V$, $\phi^V(S) \geqslant \delta_0$*

In particular, the above result implies that it is SSE -hard to certify that a graph is a vertex expander with constant expansion. This is in contrast to the case of edge expansion, where the Cheeger's inequality can be used to certify that a graph has constant edge expansion.

At the risk of being redundant, we note that our main theorem implies that any algorithm that outputs a set having vertex expansion less than $C \sqrt{\text{OPT} \log d}$ will disprove the SSE hypothesis; alternatively, to improve on the bound of $O(\sqrt{\text{OPT} \log d})$, one has to disprove the SSE hypothesis. From an algorithmic standpoint, we believe that Theorem 1.4 exposes a clean algorithmic challenge of recognizing a vertex expander – a challenging problem that is not only interesting on its own right, but whose resolution would probably lead to a significant advance in approximation algorithms.

At a high level, the proof is as follows. We introduce the notion of BALANCED ANALYTIC VERTEX EXPANSION for Markov chains. This quantity can be thought of as a CSP on $(d+1)$-tuples of vertices. We show a reduction from BALANCED ANALYTIC VERTEX EXPANSION of a Markov chain, say H, to vertex expansion of a graph, say H_1 (Section 7). Our reduction is generic and works for any Markov chain H. Surprisingly, the CSP-like nature of BALANCED ANALYTIC VERTEX EXPANSION makes it amenable to a reduction from SMALL-SET EXPANSION (Section 6). We construct a gadget for this reduction and study its embedding into the Gaussian graph to analyze its soundness (Section 4 and Section 5). The gadget involves a sampling procedure to generate a bounded-degree graph.

2   Proof Overview

BALANCED ANALYTIC VERTEX EXPANSION.   To exhibit a hardness result, we begin by defining a combinatorial optimization problem related to the problem of approximating vertex expansion in graphs having largest degree d. This problem referred to as BALANCED ANALYTIC VERTEX EXPANSION can be motivated as follows.

Fix a graph $G = (V, E)$ and a subset of vertices $S \subset V$. For any vertex $v \in V$, v is on the boundary of the set S if and only if $\max_{u \in N(v)} |\mathbb{I}_S [u] - \mathbb{I}_S [v]| = 1$, where $N(v)$ denotes the neighbourhood of vertex v. In particular, the fraction of vertices on the boundary of S is given by $\mathbb{E}_v \max_{u \in N(v)} |\mathbb{I}_S [u] - \mathbb{I}_S [v]|$. The *symmetric* vertex expansion of the set $S \subseteq V$ is given by,

$$n \cdot \frac{|N(S) \cup N(V \backslash S)|}{|S||V \backslash S|} = \frac{\mathbb{E}_v \max_{u \in N(v)} |\mathbb{I}_S [u] - \mathbb{I}_S [v]|}{\mathbb{E}_{u,v} |\mathbb{I}_S [u] - \mathbb{I}_S [v]|} .$$

Note that for a degree d graph, each of the terms in the numerator is maximization over the d edges incident at the vertex. The formal definition of BALANCED ANALYTIC VERTEX EXPANSION is as shown below.

Definition 2.1. An instance of BALANCED ANALYTIC VERTEX EXPANSION, denoted by $(V, \mathcal{P})$, consists of a set of variables V and a probability distribution $\mathcal{P}$ over $(d + 1)$-tuples in V^{d+1}. The probability distribution $\mathcal{P}$ satisfies the condition that all its $d + 1$ marginal distributions are the same (denoted by μ). The goal is to solve the following optimization problem

$$\Phi(V, \mathcal{P}) \stackrel{\text{def}}{=} \min_{F:V \mapsto \{0,1\}| \mathbb{E}_{X,Y \sim \mu}|F(X)-F(Y)| \geqslant \frac{1}{100}} \frac{\mathbb{E}_{(X,Y_1,\ldots,Y_d) \sim \mathcal{P}} \max_i |F(Y_i) - F(X)|}{\mathbb{E}_{X,Y \sim \mu} |F(X) - F(Y)|}$$

For constant d, this could be thought of as a constraint satisfaction problem (CSP) of arity $d + 1$. Every d-regular graph G has an associated instance of BALANCED ANALYTIC VERTEX EXPANSION whose value corresponds to the vertex expansion of G. Conversly, we exhibit a reduction from BALANCED ANALYTIC VERTEX EXPANSION to problem of approximating vertex expansion in a graph of degree $\mathsf{poly}(d)$ (Section 7 for details).
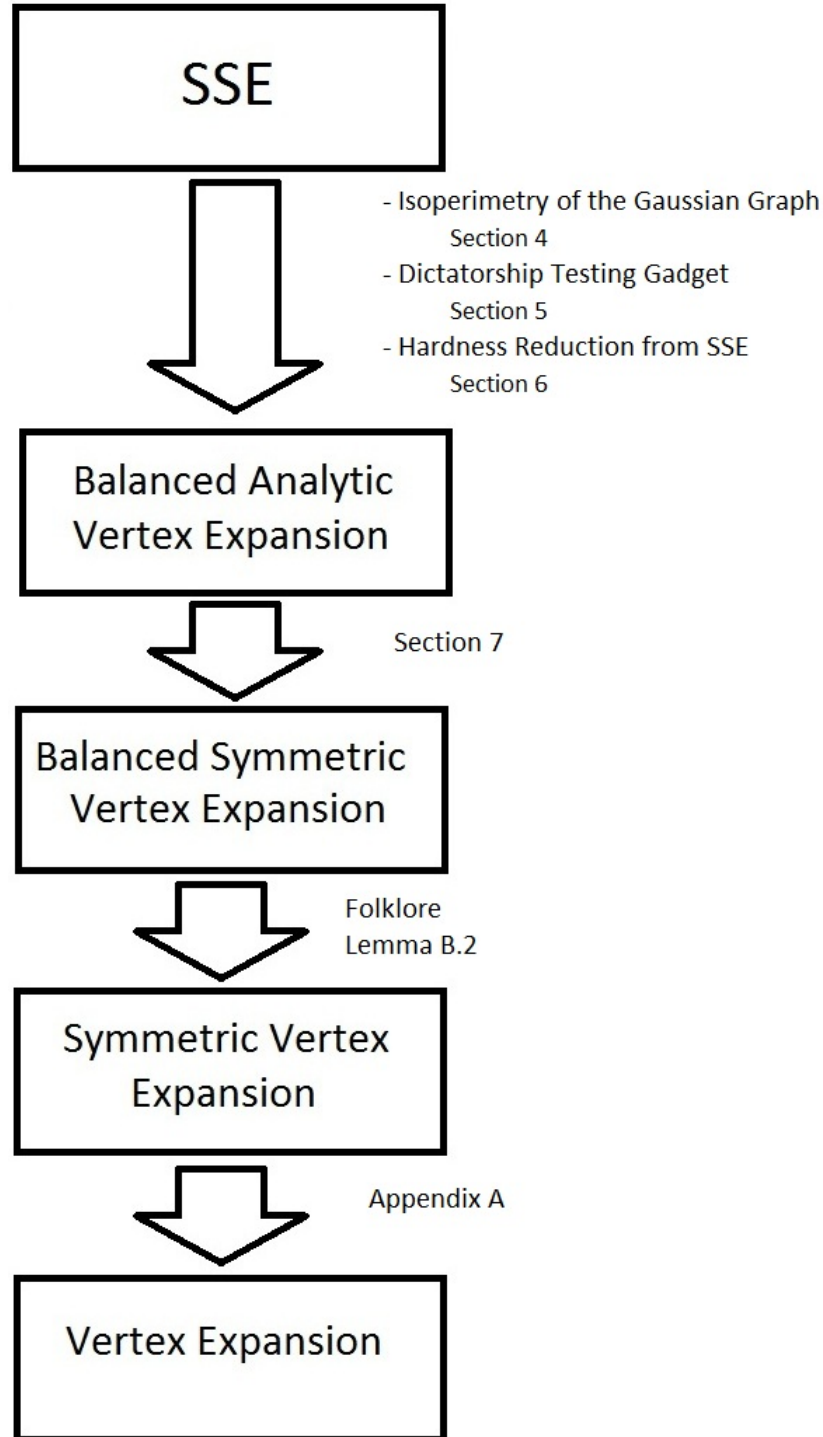
Figure 1: Reduction from SSE to Vertex Expansion

Dictatorship Testing Gadget. As with most hardness results obtained via the label cover or the unique games problem, central to our reduction is an appropriate dictatorship testing gadget.

Simply put, a dictatorship testing gadget for BALANCED ANALYTIC VERTEX EXPANSION is an instance $\mathcal{H}^R$ of the problem such that, on one hand there exists the so-called *dictator* assignments with value ε, while every assignment far from every dictator incurs a cost of at least $\Omega(\sqrt{\varepsilon \log d})$.

The construction of the dictatorship testing gadget is as follows. Let H be a Markov chain on vertices $\{s, t, t', s'\}$ connected to form a path of length three. The transition probabilities of the Markov chain $\mathcal{H}$ are so chosen to ensure that if μ_H is the stationary distribution of H then $\mu_H(t) = \mu_H(t') = \varepsilon/2$ and $\mu_H(s) = \mu_H(s') = (1 - \varepsilon)/2$. In particular, H has a vertex separator $\{t, t'\}$ whose weight under the stationary distribution is only ε.

The dictatorship testing gadget is over the product Markov chain H^R for some large constant R. The constraints $\mathcal{P}$ of the dictatorship testing gadget H^R are given by the following sampling procedure,

– Sample $x \in H^R$ from the stationary distribution of the chain.

– Sample d-neighbours $y_1, \ldots, y_d \in H^R$ of x independently from the transition probabilities of the chain H^R. Output the tuple $(x, y_1, \ldots, y_d)$.

For every $i \in [R]$, the ith dictator solution to the above described gadget is given by the following function,

$$F(x) = \begin{cases} 1 & \text{if } x_i \in \{s, t\} \\ 0 & \text{otherwise} \end{cases}$$

It is easy to see that for each constraint $(x, y_1, \ldots, y_d) \sim \mathcal{P}$, $\max_j \left| F(x) - F(y_j) \right| = 0$ unless $x_i = t$ or $x_i = t'$. Since x is sampled from the stationary distribution for μ_H, $x_i \in \{t, t'\}$ happens with probability ε. Therefore the expected cost incurred by the ith dictator assignment is at most ε.

Soundness Analysis of the Gadget. The soundness property desired of the dictatorship testing gadget can be stated in terms of influences. Specifically, given an assignment $F : V(H)^R \to [0, 1]$, the influence of the ith coordinate is given by $\mathsf{Inf}_i[F] = \mathbb{E}_{x_{[R] \setminus i}} \mathsf{Var}_{x_i}[F(x)]$, i.e., the expected variance of the function after fixing all but the ith coordinate randomly. Henceforth, we will refer to a function $F : H^R \to [0, 1]$ as *far from every dictator* if the influence of all of its coordinates are small (say $< \tau$).

We show that the dictatorship testing gadget H^R described above satisfies the following soundness – for every function F that is far from every dictator, the cost of F is at least $\Omega(\sqrt{\varepsilon \log d})$. To this end, we appeal to the invariance principle to translate the cost incurred to a corresponding isoperimetric problem on the Gaussian space. More precisely, given a function $F : H^R \to [0, 1]$, we express it as a polynomial in the eigenfunctions over H. We carefully construct a Gaussian ensemble with the same moments up to order two, as the eigenfunctions at the query points $(x, y_1, \ldots, y_d) \in \mathcal{P}$. By appealing to the invariance principle for low degree polynomials, this translates in to the following isoperimetric question over Gaussian space $\mathcal{G}$.,

Suppose we have a subset $S \subseteq \mathcal{G}$ of the n-dimensional Gaussian space. Consider the following experiment:

– Sample a point $z \in \mathcal{G}$ the Gaussian space.

– Pick d independent perturbations $z'_1, z'_2, \ldots, z'_d$ of the point z by ε-noise.

– Output 1 if at least one of the edges (z, z'_i) crosses the cut $(S, \bar{S})$ of the Gaussian space.

Among all subsets S of the Gaussian space with a given volume, which set has the least expected output in the above experiment? The answer to this isoperimetric question corresponds to the soundness of the dictatorship test. A halfspace of volume $\frac{1}{2}$ has an expected output of $\sqrt{\varepsilon \log d}$ in the above experiment. We show that among all subsets of constant volume, halfspaces acheive the least expected output value.

This isoperimetric theorem proven in Section 4 yields the desired $\Omega(\sqrt{\varepsilon \log d})$ bound for the soundness of the dictatorship test constructed via the Markov chain H. Here the noise rate of ε arises from the fact that all the eigenfunctions of the Markov chain H have an eigenvalue smaller than $1 - \varepsilon$. The details of the argument based on invariance principle is presented in Section 5

We show a $\Omega(\sqrt{\varepsilon \log d})$ lower bound for the isoperimetric problem on the Gaussian space. The proof of this isoperimetric inequality is included in Section 4

We would like to point out here that the traditional noisy cube gadget does not suffice for our application. This is because in the noisy cube gadget while the dictator solutions have an edge expansion of ε they have a vertex expansion of εd, yielding a much worse value than the soundness.

Reduction from SMALL-SET EXPANSION problem. Gadget reductions from the UNIQUE GAMES problem cannot be used towards proving a hardness result for edge or vertex expansion problems. This is because if the underlying instance of UNIQUE GAMES has a small vertex separator, then the graph produced via a gadget reduction would also have small vertex expansion. Therefore, we appeal to a reduction from the SMALL-SET EXPANSION problem (Section 6 for details).

Raghavendra et al. [RST12] show optimal inapproximability results for the Balanced separator problem using a reduction from the SMALL-SET EXPANSION problem. While the overall approach of our reduction is similar to theirs, the details are subtle. Unlike hardness reductions from unique games, the reductions for expansion-type problems starting from SMALL-SET EXPANSION are not very well understood. For instance, the work of Raghavendra and Tan [RT12] gives a dictatorship testing gadget for the Max-Bisection problem, but a SMALL-SET EXPANSION based hardness for Max-Bisection still remains open.

2.1 Notation

We use μ_G to denote a probability distribution on vertices of the graph G. We drop the subscript G, when the graph is clear from the context. For a set of vertices S, we define $\mu(S) = \int_{x \in S} \mu(x)$. We use $\mu_{|S}$ to denote the distribution μ restricted to the set $S \subset V(G)$. If the graph has vertex weights, we denote the weight of a vertex $v \in V(G)$ by $w(v)$ and denote the weight of a set $S \subseteq V$ by $w(S)$. We denote the degree of a vertex v by $\deg(v)$. We denote the neighborhood of S in G by $N_G(S)$, i.e.

$$N_G(S) = \{v \in \bar{S} \,|\, \exists u \in S \text{ such that } \{u, v\} \in E(G)\} \,.$$

We drop the subscript G when the graph is clear from the context.

2.2 Organization

We begin with some definitions and the statements of the SSE hypotheses in Section 3. In Section A, we show that the computation of vertex expansion and symmetric vertex expansion is equivalent upto constant factors. We prove a new Gaussian isoperimetry results in Section 4 that we use in our soundness analysis. In Section 5 we show the construction of our main gadget and analyze its soundness and completeness using BALANCED ANALYTIC VERTEX EXPANSION as the test function. We show a reduction from a reduction from BALANCED ANALYTIC VERTEX EXPANSION to vertex expansion in Section 7. In Section 6, we use this gadget to

show a reduction SSE to BALANCED ANALYTIC VERTEX EXPANSION. Finally, in Section 8, we show how to put all the reductions togethor to get optimal SSE -hardness for vertex expansion.

Complimenting our lower bound, we give an algorithm that outputs a set having vertex expansion at most $O(\sqrt{\mathsf{OPT} \log d})$ in Section 9.

3 Preliminaries

Symmetric Vertex Expansion. For our proofs, the notion of Symmetric Vertex Expansion is useful.

Definition 3.1. Given a graph $G = (V, E)$, we define the the symmetric vertex expansion of a set $S \subset V$ as follows.

$$\Phi_G^{\mathsf{V}}(S) \stackrel{\text{def}}{=} n \cdot \frac{|N_G(S) \cup N_G(V \setminus S)|}{|S||V \setminus S|}$$

Balanced Vertex Expansion. We define the balanced vertex expansion of a graph as follows.

Definition 3.2. Given a graph G and balance parameter b, we define the *b-balanced vertex expansion* of G as follows.

$$\phi_b^{\mathsf{V},\mathsf{bal}} \stackrel{\text{def}}{=} \min_{S : |S||V \setminus S| \geqslant bn^2} \phi^{\mathsf{V}}(S).$$

and

$$\Phi_b^{\mathsf{V},\mathsf{bal}} \stackrel{\text{def}}{=} \min_{S : |S||V \setminus S| \geqslant bn^2} \Phi^{\mathsf{V}}(S).$$

We define $\phi^{\mathsf{V},\mathsf{bal}} \stackrel{\text{def}}{=} \phi_{1/100}^{\mathsf{V},\mathsf{bal}}$ and $\Phi^{\mathsf{V},\mathsf{bal}} \stackrel{\text{def}}{=} \Phi_{1/100}^{\mathsf{V},\mathsf{bal}}$.

Analytic Vertex Expansion. Our reduction from SSE to vertex expansion goes via an intermediate problem that we call d-BALANCED ANALYTIC VERTEX EXPANSION. We define the notion of d-BALANCED ANALYTIC VERTEX EXPANSION as follows.

Definition 3.3. An instance of d-BALANCED ANALYTIC VERTEX EXPANSION, denoted by $(V, \mathcal{P})$, consists of a set of variables V and a probability distribution $\mathcal{P}$ over $(d + 1)$-tuples in V^{d+1}. The probability distribution $\mathcal{P}$ satisfies the condition that all its $d + 1$ marginal distributions are the same (denoted by μ). The d-BALANCED ANALYTIC VERTEX EXPANSION under a function $F : V \mapsto \{0, 1\}$ is defined as

$$\Phi(V, \mathcal{P})(F) \stackrel{\text{def}}{=} \frac{\mathbb{E}_{(X,Y_1,\ldots,Y_d) \sim \mathcal{P}} \max_i |F(Y_i) - F(X)|}{\mathbb{E}_{X,Y \sim \mu} |F(X) - F(Y)|}.$$

The d-BALANCED ANALYTIC VERTEX EXPANSION of $(V, \mathcal{P})$ is defined as

$$\Phi(V, \mathcal{P}) \stackrel{\text{def}}{=} \min_{F : V \mapsto \{0,1\} | \mathbb{E}_{X,Y \sim \mu} |F(X) - F(Y)| \geqslant \frac{1}{100}} \Phi(V, \mathcal{P})(F).$$

When drop the degree d from the notation, when it is clear from the context.

For an instance $(V, \mathcal{P})$ of BALANCED ANALYTIC VERTEX EXPANSION and an assignment $F : V \mapsto \{0, 1\}$ define

$$\mathrm{val}_{\mathcal{P}}(F) = \mathbb{E}_{(X,Y_1,\ldots,Y_d) \sim \mathcal{P}} \max_i |F(Y_i) - F(X)|.$$

Gaussian Graph. Recall that two standard normal random variables X, Y are said to be α-correlated if there exists an independent standard normal random variable Z such that $Y = \alpha X + \sqrt{1 - \alpha^2} Z$.

Definition 3.4. The *Gaussian Graph* $\mathcal{G}_{\Lambda,\Sigma}$ is a complete weighted graph on the vertex set $V(\mathcal{G}_{\Lambda,\Sigma}) = \mathbb{R}^n$. The weight of the edge between two vertices $u, v \in V(\mathcal{G}_{\Lambda,\Sigma})$ is given by

$$w(\{u, v\}) = \mathbb{P}[X = u \text{ and } Y = v]$$

where $Y \sim \mathcal{N}(\Lambda X, \Sigma)$, where Λ is a diagonal matrix such that $\|\Lambda\| \leqslant 1$ and $\Sigma \geq \varepsilon I$ is a diagonal matrix.

Remark 3.5. Note that for any two non-empty disjoint sets $S_1, S_2 \subset V(\mathcal{G}_{\Lambda,\Sigma})$, the total weight of the edges between S_1 and S_2 can be non-zero even though every single edge in the $\mathcal{G}_{\Lambda,\Sigma}$ has weight zero.

Definition 3.6. We say that a graph $G = (V, E)$ is $\Theta(d)$-regular, if there exist absolute constants $c_1, c_2 \in \mathbb{R}^+$ such that for all vertices $i \in V$, we have $c_1 d \leqslant \deg(i) \leqslant c_2 d$.

For a random variable X, define the variance and ℓ_1-variance as follows,

$$\mathsf{Var}[X] = \underset{X_1, X_2}{\mathbb{E}} [(X_1 - X_2)^2] \qquad \mathsf{Var}_1[X] = \underset{X_1, X_2}{\mathbb{E}} [|X_1 - X_2|]$$

where X_1, X_2 are two independent samples of X.

Small-Set Expansion Hypothesis.

Problem 3.7 (SMALL-SET EXPANSION (γ, δ))**.** Given a regular graph $G = (V, E)$, distinguish between the following two cases:

Yes: There exists a non-expanding set $S \subset V$ with $\mu(S) = \delta$ and $\Phi_G(S) \leqslant \gamma$.

No: All sets $S \subset V$ with $\mu(S) = \delta$ are highly expanding having $\Phi_G(S) \geqslant 1 - \gamma$.

Hypothesis 3.8 (Hardness of approximating SMALL-SET EXPANSION)**.** *For all $\gamma > 0$, there exists $\delta > 0$ such that the promise problem* SMALL-SET EXPANSION *(γ, δ) is* NP-*hard.*

For the proofs, it shall be more convenient to use the following version of the SMALL-SET EXPANSION problem, in which we high expansion is guaranteed not only for sets of measure δ, but also within an arbitrary multiplicative factor of δ.

Problem 3.9 (SMALL-SET EXPANSION (γ, δ, M))**.** Given a regular graph $G = (V, E)$, distinguish between the following two cases:

Yes: There exists a non-expanding set $S \subset V$ with $\mu(S) = \delta$ and $\Phi_G(S) \leqslant \gamma$.

No: All sets $S \subset V$ with $\mu(S) \in \left(\frac{\delta}{M}, M\delta \right)$ have $\Phi_G(S) \geqslant 1 - \gamma$.

The following stronger hypothesis was shown to be equivalent to Small-Set Expansion Hypothesis in [RST12].

Hypothesis 3.10 (Hardness of approximating SMALL-SET EXPANSION)**.** *For all $\gamma > 0$ and $M \geqslant 1$, there exists $\delta > 0$ such that the promise problem* SMALL-SET EXPANSION *(γ, δ, M) is* NP-*hard.*

4 Isoperimetry of the Gaussian Graph

In this section we bound the Balanced Analytic Vertex Expansion of the Gaussian graph. For the Gaussian Graph, we define the canonical probability distribution on V^{d+1} as follows. The marginal distribution along any component X or Y_i is the standard Gaussian distribution in $\mathbb{R}^n$, denoted here by $\mu = \mathcal{N}(0,1)^n$.

$$\mathcal{P}_{\mathcal{G}_{\Lambda,\Sigma}}(X, Y_1, \ldots, Y_d) = \frac{\Pi_{i=1}^d w(X, Y_i)}{\mu(X)^{d-1}} = \mu(X)\Pi_{i=1}^d \mathbb{P}[Y = Y_i].$$

Here, random variable Y is sampled from $\mathcal{N}(\Lambda X, \Sigma)$.

Theorem 4.1. *For any closed set $S \subset of V(\mathcal{G}_{\Lambda,\Sigma})$ ($\Sigma \geq \varepsilon I$ is a diagonal matrix), we have*

$$\frac{\mathbb{E}_{(X,Y_1,\ldots,Y_d)\sim\mathcal{P}_{\mathcal{G}_{\Lambda,\Sigma}}} \max_i |\mathbb{I}_S[X] - \mathbb{I}_S[Y_i]|}{\mathbb{E}_{X,Y\sim\mu}|\mathbb{I}_S[X] - \mathbb{I}_S[Y]|} = \frac{\mathbb{E}_{X\sim\mu}\mathbb{E}_{Y_1,\ldots Y_d\sim\mathcal{N}(\Lambda X,\Sigma)} \max_i |\mathbb{I}_S[X] - \mathbb{I}_S[Y_i]|}{\mathbb{E}_{X,Y\sim\mu}|\mathbb{I}_S[X] - \mathbb{I}_S[Y]|} \geq c\sqrt{\varepsilon\log d}$$

for some absolute constant c.

Lemma 4.2. *Let $u, v \in \mathbb{R}^n$ satisfy $|u - v| \leq \sqrt{\varepsilon\log d}$ and let Σ be a diagonal matrix satisfying $\Sigma \geq \varepsilon I$. Let P_u, P_v be the distributions $\mathcal{N}(\Lambda u, \Sigma)$ and $\mathcal{N}(\Lambda v, \Sigma)$ respectively. Then,*

$$d_{\mathsf{TV}}(P_u, P_v) \leq 1 - \frac{1}{d}.$$

Proof. First, we note that that for the purpose of estimating their total variation distance, we can view P_u, P_v as one-dimensional Gaussians along the line $\Lambda u - \Lambda v$. Since $\|\Lambda\| \leq 1$,

$$\|\Lambda u - \Lambda v\| \leq \|u - v\| \leq \sqrt{\varepsilon\log d}.$$

Wlog, we may take $\Lambda u = 0$ and $\Lambda v = \sqrt{\varepsilon\log d}$. Next, by the definition of total variation distance,

$$
\begin{aligned}
d_{\mathsf{TV}}(P_u, P_v) &= \int_{x:P_v(x)\geq P_u(x)} |P_v(x) - P_u(x)|dx \\
&= \int_{\Lambda v/2}^{\infty} (P_v(x) - P_u(x))dx \\
&= \frac{1}{\sqrt{2\pi\varepsilon}}\int_{\Lambda v/2}^{\infty} e^{-\frac{\|x-\Lambda v\|^2}{2\varepsilon}}dx - \frac{1}{\sqrt{2\pi\varepsilon}}\int_{\Lambda v/2}^{\infty} e^{-\frac{\|x\|^2}{2\varepsilon}}dx \\
&= \frac{1}{\sqrt{2\pi\varepsilon}}\int_{-\Lambda v/2}^{\Lambda v/2} e^{-\frac{\|x\|^2}{2\varepsilon}}dx \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\sqrt{\log d}/2}^{\sqrt{\log d}/2} e^{-\frac{\|x\|^2}{2}}dx \\
&= 1 - 2\cdot\frac{1}{\sqrt{2\pi}}\int_{\sqrt{\log d}/2}^{\infty} e^{-\frac{\|x\|^2}{2}}dx \\
&< 1 - \frac{1}{d}.
\end{aligned}
$$

where the last step uses a standard bound on the Gaussian tail.

$\square$

Proof of Theorem 4.1. Let μ_X denote the Gaussian distribution $\mathcal{N}(\Lambda X, \Sigma)$. Then the LHS is:

$$\int_{\mathbb{R}^n \setminus S} \left(1 - (1 - \mu_X(S))^d\right) d\mu(X) + \int_S \left(1 - (1 - \mu_X(\mathbb{R}^n \setminus S))^d\right) d\mu(X).$$

To bound this, we will restrict ourselves to points X for which the μ_X measure of the complementary set is at least $1/d$. Roughly speaking, these will be points near the boundary of S. Define:

$$S_1 = \left\{ x \in S \ : \ \mu_X(\mathbb{R}^n \setminus S) < \frac{1}{2d} \right\}, \ S_2 = \left\{ x \in \mathbb{R}^n \setminus S \ : \ \mu_X(S) < \frac{1}{2d} \right\}$$

and

$$S_3 = \mathbb{R}^n \setminus S_1 \setminus S_2.$$

For $u \in \mathbb{R}^n$, let P_u be the distribution $\mathcal{N}(\Lambda u, \Sigma)$. For any $u \in S_1, v \in S_2$, we have

$$d_{\mathsf{TV}}(P_u, P_v) > 1 - \frac{1}{2d} - \frac{1}{2d} = 1 - \frac{1}{d}.$$

Therefore, by Lemma 4.2, $\|u - v\| > \sqrt{\varepsilon \log d}$, i.e., $d(S_1, S_2) > \sqrt{\varepsilon \log d}$. Next we bound the measure of S_3. We can assume wlog that $\mu(S) \leqslant \mu(\mathbb{R}^n \setminus S)$ and $\mu(S_1) \geqslant \mu(S)/2$ (else $\mu(S_3) \geqslant \mu(S)/2$ and we are done). Applying the isoperimetric inequality for Gaussian space (see surveys [Vem05, VLM10]), for subsets at this distance,

$$\mu(S_3) \geqslant \sqrt{\frac{2}{\pi}} \sqrt{\varepsilon \log d} \cdot \mu(S_1)\mu(S_2) \geqslant \sqrt{\frac{\varepsilon \log d}{2\pi}} \cdot \mu(S)\mu(\mathbb{R}^n \setminus S).$$

We are now ready to complete the proof.

$$\frac{1}{2} \left(\int_{\mathbb{R}^n \setminus S} (1 - (1 - \mu_X(S))^d) d\mu(X) + \int_S (1 - (1 - \mu_X(\mathbb{R}^n \setminus S))) d\mu(X) \right)$$

$$\geqslant \frac{1}{2} \left(\int_{X \in \mathbb{R}^n \setminus S, \mu_X(S) \geqslant 1/d} (1 - (1 - \mu_X(S))^d) d\mu(X) + \int_{X \in S, \mu_X(\mathbb{R}^n \setminus S) \geqslant 1/d} (1 - (1 - \mu_X(\mathbb{R}^n \setminus S))) d\mu(X) \right)$$

$$\geqslant \frac{e - 1}{2e} \left(\int_{X \in \mathbb{R}^n \setminus S, \mu_X(S) \geqslant 1/d} d\mu(X) + \int_{X \in S, \mu_X(\mathbb{R}^n \setminus X) \geqslant 1/d} d\mu(X) \right)$$

$$\geqslant \frac{e - 1}{2e} \mu(S_3)$$

$$\geqslant c \sqrt{\varepsilon \log d} \cdot \mu(S)\mu(\mathbb{R}^n \setminus S).$$

$\square$

We prove the following Theorem which helps us to bound the isoperimetry of the Gaussian graph for over all functions over the range $[0, 1]$.

Theorem 4.3. *Given an instance $(V, \mathcal{P})$ and a function $F : V \mapsto [0, 1]$, there exists a function $F' : V \mapsto \{0, 1\}$, such that*

$$\frac{\mathbb{E}_{(X, Y_1, \ldots, Y_d) \sim \mathcal{P}} \max_i |F(X) - F(Y_i)|}{\mathbb{E}_{X, Y \sim \mu} |F(X) - F(Y)|} \geqslant \frac{\mathbb{E}_{(X, Y_1, \ldots, Y_d) \sim \mathcal{P}} \max_i |F'(X) - F'(Y_i)|}{\mathbb{E}_{X, Y \sim \mu} |F'(X) - F'(Y)|}$$

Proof. For every $r \in [0, 1]$, we define $F_r : V \mapsto \{0, 1\}$ as follows.

$$F_r(X) = \begin{cases} 1 & F(X) \geqslant r \\ 0 & F(X) < r \end{cases}$$

Clearly,

$$F(X) = \int_0^1 F_r(X) dr.$$

Now, observe that if $F(X) - F(Y) \geqslant 0$ then $F_r(X) - F_r(Y) \geqslant 0 \ \forall r \in [0, 1]$ and similiarly, if $F(X) - F(Y) < 0$ then $F_r(X) - F_r(Y) \leqslant 0 \ \forall r \in [0, 1]$. Therefore,

$$|F(X) - F(Y)| = \left| \int_0^1 (F_r(X) - F_r(Y)) \, dr \right| = \int_0^1 |F_r(X) - F_r(Y)| \, dr.$$

Also, observe that if $|F(X) - F(Y_1)| \geqslant |F(Y_i) - F(X)|$ then

$$|F_r(X) - F_r(Y_1)| \geqslant |F_r(Y_i) - F_r(X)| \ \ \forall r \in [0, 1]$$

Therefore,

$$
\begin{aligned}
\frac{\mathbb{E}_{(X,Y_1,\ldots,Y_d)\sim\mathcal{P}} \max_i |F(X) - F(Y_i)|}{\mathbb{E}_{X,Y\sim\mu} |F(X) - F(Y)|} \ &= \ \frac{\mathbb{E}_{(X,Y_1,\ldots,Y_d)\sim\mathcal{P}} \max_i \int_0^1 |F_r(X) - F_r(Y_i)| \, dr}{\mathbb{E}_{X,Y\sim\mu} \int_0^1 |F_r(X) - F_r(Y)| \, dr} \\
&= \ \frac{\int_0^1 \left(\mathbb{E}_{(X,Y_1,\ldots,Y_d)\sim\mathcal{P}} \max_i |F_r(X) - F_r(Y_i)| \right) dr}{\int_0^1 \left(\mathbb{E}_{X,Y\sim\mu} |F_r(X) - F_r(Y)| \right) dr} \\
&\geqslant \ \min_{r\in[0,1]} \frac{\mathbb{E}_{(X,Y_1,\ldots,Y_d)\sim\mathcal{P}} \max_i |F_r(X) - F_r(Y_i)|}{\mathbb{E}_{X,Y\sim\mu} |F_r(X) - F_r(Y)|}
\end{aligned}
$$

Let r' be the value of r which minimizes the expression above. Taking F' to be $F_{r'}$ finishes the proof. □

Corollary 4.4 (Corollary to Theorem 4.1 and Theorem 4.3)**.** *Let* $F : V(\mathcal{G}_{\Lambda,\Sigma}) \mapsto [0, 1]$ *be any function. Then, for some absolute constant c,*

$$\frac{\mathbb{E}_{(X,Y_1,\ldots,Y_d)\sim\mathcal{P}_{\mathcal{G}_{\Lambda,\Sigma}}} \max_i |F(X) - F(Y_i)|}{\mathbb{E}_{X,Y\sim\mu} |F(X) - F(Y)|} \geqslant c \sqrt{\varepsilon \log d}.$$

5 Dictatorship Testing Gadget

In this section we initiate the construction of the dictatorship testing gadget for reduction from SSE .

Overall, the dictatorship testing gadget is obtained by picking an appropriately chosen constant sized Markov-chain H, and considering the product Markov chain H^R. Formally, given a Markov chain H, define an instance of Balanced Analytic Vertex Expansion with vertices as V_H and the constraints given by the following canonical probability distribution over V_H^{d+1}.

- Sample $X \sim \mu_H$, the stationary distribution of the Markov chain V_H.

– Sample $Y_1, \ldots, Y_d$ independently from the neighbours of X in V_H

For our application, we use a specific Markov chain H on four vertices. Define a Markov chain H on $V_H = \{s, t, t', s'\}$ as follows, $p(s|s) = p(s'|s') = 1 - \frac{\varepsilon}{1-2\varepsilon}$, $p(t|s) = p(t'|s') = \frac{\varepsilon}{1-2\varepsilon}$, $p(s|t) = p(s'|t') = \frac{1}{2}$ and $p(t'|t) = p(t|t') = \frac{1}{2}$. It is easy to see that the stationary distribution of the Markov chain H over V_H is given by,

$$\mu_H(s) = \mu_H(s') = \frac{1}{2} - \varepsilon \qquad \mu_H(t) = \mu_H(t') = \varepsilon$$

From this Markov chain, construct a dictatorship testing gadget $(V_H^R, \mathcal{P}_H^R)$ as described above. We begin by showing that this dictatorship testing gadget has small vertex separators corresponding to dictator functions.

Proposition 5.1 (Completeness). *For each $i \in [R]$, the ith-dictator set defined as $F(x) = 1$ if $x_i \in \{s, t\}$ and 0 otherwise satisfies,*

$$\mathsf{Var}_1[F] = \frac{1}{2} \qquad and \qquad \mathrm{val}_{\mathcal{P}_{H^R}}(F) \leqslant 2\varepsilon$$

Proof. Clearly,

$$\mathop{\mathbb{E}}_{X,Y \sim \mu_H} |F(X) - F(Y)| = \frac{1}{2}$$

Observe that for any choice of $(X, Y_1, \ldots, Y_d) \sim \mathcal{P}_{H^R}$, $\max_i |F(X) - F(Y_i)|$ is non-zero if and only if either $x_i = t$ or $x_i = t'$. Therefore we have,

$$\mathop{\mathbb{E}}_{(X,Y_1,\ldots,Y_d) \sim \mathcal{P}_H} \max_i |F(X) - F(Y_i)| \leqslant \mathbb{P}[x_i \in \{t, t'\}]) = 2\varepsilon,$$

which concludes the proof. $\qquad\square$

5.1 Soundness

We will show a general soundness claim that holds for dictatorship testing gadgets $(V(H^R), \mathcal{P}_{H^R})$ constructed out of arbitrary Markov chains H with a given spectral gap. Towards formally stating the soundness claim, we recall some background and notation about polynomials over the product Markov chain H^R.

5.2 Polynomials over H^R

In this section, we recall how functions over the product Markov chain H^R can be written as multilinear polynomials over the eigenfunctions of H.

Let $e_0, e_1, \ldots, e_n : V(H) \to \mathbb{R}$ be an orthonormal basis of eigenvectors of H and let $\lambda_0, \ldots, \lambda_n$ be the corresponding eigenvalues. Here $e_0 = 1$ is the constant function whose eigenvalue $\lambda_0 = 1$. Clearly $e_0, \ldots, e_n$ form an orthonormal basis for the vector space of functions from $V(H)$ to $\mathbb{R}$.

It is easy to see that the eigenvectors of the product chain H^R are given by products of $e_0, \ldots, e_n$. Specifically, the eigenvectors of H^R are indexed by $\sigma \in [n]^R$ as follows,

$$e_\sigma(x) = \prod_{i=1}^{R} e_{\sigma_i}(x_i)$$

Every function $f : H^R \to \mathbb{R}$ can be written in this orthonormal basis $f(x) = \sum_{\sigma \in [n]^R} \hat{f}_\sigma e_\sigma(x)$. For a multi-index $\sigma \in [n]^R$, the function e_σ is a monomial of degree $|\sigma| = |\{i | \sigma_i \neq 0\}|$.

For a polynomial $Q = \sum_\sigma \hat{Q}_\sigma e_\sigma$, the polynomial $Q^{>p}$ denotes the projection on to degrees higher than p, i.e., $Q^{>p} = \sum_{\sigma, |\sigma|>p} \hat{Q}_\sigma e_\sigma$. The influences of a polynomial $Q = \sum_\sigma \hat{Q}_\sigma$ are defined as,

$$\mathsf{Inf}_i(Q) = \sum_{\sigma:\sigma_i \neq 0} \hat{Q}_\sigma^2$$

The above notions can be naturally extended to vectors of multilinear polynomials $Q = (Q_0, Q_1, \ldots, Q_d)$.

Note that every real-valued function on the vertices $V(H)$ of a Markov chain H can be thought of as a random variable. For each $i > 0$, the random variable $e_i(x)$ has mean zero and variance 1. The same holds for all $e_\sigma(x)$ for all $|\sigma| \neq 0$. For a function $Q : V(H^R) \to \mathbb{R}$ (or equivalently a polynomial), $\mathsf{Var}[Q]$ denotes the variance of the random variable $Q(x)$ for a random x from stationary distribution of H^R. It is an easy computation to check that this is given by,

$$\mathsf{Var}[Q] = \sum_{\sigma:|\sigma| \neq 0} \hat{Q}_\sigma^2$$

We will make use of the following Invariance Principle due to Isaksson and Mossel [IM12].

Theorem 5.2 ([IM12])**.** *Let $X = (X_1, \ldots, X_n)$ be an independent sequence of ensembles, such that $\mathbb{P}[X_i = x] \geqslant \alpha > 0, \forall i, x$. Let Q be a d-dimensional multilinear polynomial such that $\mathsf{Var}(Q_j(X)) \leqslant 1$, $\mathsf{Var}(Q_j^{>p}) \leqslant (1 - \varepsilon\eta)^{2p}$ and $\mathsf{Inf}_i(Q_j) \leqslant \tau$ where $p = \frac{1}{18} \log(1/\tau)/\log(1/\alpha)$. Finally, let $\psi : \mathbb{R}^k \mapsto \mathbb{R}$ be Lipschitz continuous. Then,*

$$\left| \mathbb{E}[\psi(Q(X))] - \mathbb{E}[\psi(Q(Z))] \right| = O(\tau^{\frac{\varepsilon\eta}{18}/\log \frac{1}{\alpha}})$$

where Z is an independent sequence of Gaussian ensembles with the same covariance structure as X.

5.3   Noise Operator

We define a noise operator $\Gamma_{1-\eta}$ on functions on the Markov chain H as follows :

$$\Gamma_{1-\eta} F(X) \stackrel{\text{def}}{=} (1 - \eta)F(X) + \eta \mathop{\mathbb{E}}_{Y \sim X} F(Y)$$

for every function $F : H \to \mathbb{R}$. Similarly, one can define the noise operator $\Gamma_{1-\eta}$ on functions over H^R.

Applying the noise operator $\Gamma_{1-\eta}$ on a function F, smoothens the function or makes it closer to a low-degree polynomial. This resulting function $\Gamma_{1-\eta}F$ is close to a *low-degree polynomial*, and therefore is amenable to applying an invariance principle. Formally, one can show the following decay of coefficients of high degree for $\Gamma_{1-\eta}F$. We defer the proof to the Appendix (Lemma C.1).

Lemma 5.3. *(Decay of High degree Coefficients) Let Q_j be the multi-linear polynomial representation of $\Gamma_{1-\eta}F(X)$, and let ε be the spectral gap of the Markov chain H. Then,*

$$\mathsf{Var}(Q_j^{>p}) \leqslant (1 - \varepsilon\eta)^{2p}$$

Furthermore, on applying the noise operator $\Gamma_{1-\eta}$, the resulting function $\Gamma_{1-\eta}F$ can have a bounded number of influential coordinats as shown by the following lemma.

Lemma 5.4. *(Sum of Influences Lemma) If the spectral gap of a Markov chain is at least ε then for any function $F : V_H^R \to \mathbb{R}$,*

$$\sum_{i \in [R]} \mathsf{Inf}_i(\Gamma_{1-\eta}F) \leqslant \frac{1}{\eta\varepsilon} \mathsf{Var}[F]$$

14

Proof. By suitable normalization, we may assume without loss of generality that $\mathsf{Var}[F] = 1$. If Q denotes the multilinear representation of $\Gamma_{1-\eta}F$, then the sum of influences can be written as,

$$
\begin{aligned}
\sum_{i \in [R]} \mathsf{Inf}_i(\Gamma_{1-\eta}F) &\leqslant \sum_{|\sigma| \neq 0} |\sigma| \hat{Q}_\sigma^2 \\
&\leqslant \sum_{|\sigma| \neq 0} |\sigma| (1 - \eta\varepsilon)^{2|\sigma|} \hat{F}_\sigma^2 \\
&\leqslant \left(\max_{k \in \mathbb{N}} k(1 - \eta\varepsilon)^{2k} \right) \sum_{|\sigma| \neq 0} \hat{F}_\sigma^2 < \frac{1}{\eta\varepsilon}
\end{aligned}
$$

where we used the fact that the function $h(t) = t(1 - \eta\varepsilon)^{2t}$ achieves its maximum value at $t = -\frac{1}{2}\ln(1 - \eta\varepsilon)$. $\quad\square$

5.4 Soundness Claim

Now we are ready to formally state our soundness claim for a dictatorship test gadget constructed out of a Markov chain.

Proposition 5.5 (Soundness). *For all $\varepsilon, \eta, \alpha, \tau > 0$ the following holds. Let H be a finite Markov-chain with a spectral gap of at least ε, and the probability of every state under stationary distribution is $\geqslant \alpha$. Let $F : V(H^R) \mapsto \{0, 1\}$ be a function such that $\max_{i \in [R]} \mathsf{Inf}_i(\Gamma_{1-\eta}F) \leqslant \tau$. Then we have*

$$
\mathop{\mathbb{E}}_{(X,Y_1,\dots,Y_d) \sim \mathcal{P}_{H^R}} [\max_i |F(Y_i) - F(X)|] \geqslant \Omega(\sqrt{\varepsilon \log d}) \mathop{\mathbb{E}}_{X,Y \sim \mu_{H^R}} |F(X) - F(Y)| - O(\eta) - \tau^{\Omega(\varepsilon\eta/\log(1/\alpha))}
$$

For the sake of brevity, we define $\mathsf{soundness}(V(H^R), \mathcal{P}_{H^R})$ to be the following :

Definition 5.6.

$$
\mathsf{soundness}(V(H^R), \mathcal{P}_{H^R}) \overset{\text{def}}{=} \min_{F:\max_{i \in [R]} \mathsf{Inf}_i(F) \leqslant \tau} \frac{\mathbb{E}_{(X,Y_1,\dots,Y_d) \sim \mathcal{P}_{H^R}}[\max_i |F(Y_i) - F(X)|]}{\mathbb{E}_{X,Y \sim \mu_{H^R}} |F(X) - F(Y)|}
$$

In the rest of the section, we will present a proof of Proposition 5.5. First, we construct gaussian random variables with moments matching the eigenvectors of the chain H.

Gaussian Ensembles. Let $Q = (Q_0, Q_1, \dots, Q_d)$ be the multi-linear polynomial representation of the vector-valued function $\left(\Gamma_{1-\eta}F(X), \Gamma_{1-\eta}F(Y_1), \dots, \Gamma_{1-\eta}F(Y_d)\right)$. Let E denote the ensemble of nd random variables $(e_0(X), e_1(X), \dots, e_n(X)), (e_0(Y_1), \dots, e_n(Y_1)), \dots, (e_0(Y_d), \dots, e_n(Y_d))$. Let $E_1, \dots, E_R$ be R independent copies of the ensemble E. Clearly, the polynomial Q can be thought of as a polynomial over $E_1, \dots, E_R$. For each random variable x in $E_1, \dots, E_R$ and a value β in its support, $\mathbb{P}[x = \beta]$ is at least the minimum probability of a vertex in H under its stationary distribution.

This polynomial Q satisfies the requirements of Theorem 5.2 because on the one hand, the influences of F are $\leqslant \tau$ and on the other by Lemma 5.3, $\mathsf{Var}(Q^{\geqslant p}) \leqslant (1 - \varepsilon\eta)^{2p}$. Now we will apply the invariance principle to relate the soundness to the corresponding quantity on the gaussian graph, and then appeal to the isoperimetric result on the Gaussian graph (Theorem 4.1).

The invariance principle translates the polynomial $(Q_0(X), Q_1(Y_1), \dots Q_d(Y_d))$ on the sequence of independent ensembles $E_1, \dots, E_R$, to a polynomial on a corresponding sequence of gaussian ensembles with the same moments up to degree two.

Consider the ensemble E. For each $i \neq 0$, the expectation $\mathbb{E}[e_i(X)] = \mathbb{E}[e_i(Y_1) = 0] = \ldots \mathbb{E}[e_i(Y_d)] = 0$. For each $i \neq j$, it is easy to see that, $\mathbb{E}[e_i(X)e_j(X)] = \mathbb{E}[e_i(Y_1)e_j(Y_1)] = \ldots \mathbb{E}[e_i(Y_d)e_j(Y_d)] = 0$. Moreover, $\mathbb{E}[e_i(X)e_j(Y_a)] = \mathbb{E}[e_i(Y_a)e_j(Y_b)] = 0$ whenever $i \neq j$ and all $a, b \in \{1, \ldots d\}$. The only non-trivial correlations are $\mathbb{E}[e_i(X)e_i(Y_a)]$ and $\mathbb{E}[e_i(Y_a)e_i(Y_b)]$ for all $i \in [n]$ and $a, b \in [d]$. It is easy to check that

$$\mathbb{E}[e_i(X)e_i(Y_a)] = \lambda_i \qquad\qquad \mathbb{E}[e_i(Y_a)e_i(Y_b)] = \lambda_i^2$$

From the above discussion, we see that the following gaussian ensemble $z = (z_X, z_{Y_1}, \ldots, z_{Y_d})$ has the same covariance as the ensemble E.

1. Sample z_X and n-dimensional Gaussian random vector.

2. Sample $z_{Y_1}, \ldots, z_{Y_d} \in \mathbb{R}^n$ i.i.d as follows : The ith coordinate of each z_{Y_a} is sampled from $\lambda_i z_X(i) + \sqrt{1 - \lambda_i^2}\xi_{a,i}$ where $\xi_{a,i}$ is a Gaussian random variable independent of z_X and all other $\xi_{a,i}$.

Let $Z_X, Z_{Y_1}, \ldots, Z_{Y_d} \in \mathbb{R}^{nR}$ be the ensemble obtained by R independent samples from $z_X, z_{Y_1}, \ldots, z_{Y_d}$.

Let Σ denote the $nR \times nR$ diagonal matrix whose entries are $1 - \lambda_1^2, \ldots, 1 - \lambda_n^2$ repeated R times. Since the spectral gap of H is ε, we have that $1 - \lambda_i^2 \geqslant 2\varepsilon - \varepsilon^2 > \varepsilon$ for all $i \in \{1, \ldots, n\}$. Therefore, we have $\Sigma > \varepsilon I$.

Proof of soundness. Now we return to the proof of the main soundness claim for the dictatorship testing gadget $(V(H^R), \mathcal{P}_{\mathcal{H}^R})$ constructed out an arbitrary Markov chain.

Proof of Proposition 5.5. Let $Q = (Q_0, Q_1, \ldots, Q_d)$ be the multi-linear polynomial representation of the vector-valued function $\left(\Gamma_{1-\eta}F(X), \Gamma_{1-\eta}F(Y_1), \ldots, \Gamma_{1-\eta}F(Y_d)\right)$.

Define a function $s : \mathbb{R} \to \mathbb{R}$ as follows

$$s(x) = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } x \in [0, 1] \\ 1 & \text{if } x > 1 \end{cases}$$

Define a function $\Psi : \mathbb{R}^{d+1} \to \mathbb{R}$ as, $\Psi(x, y_1, \ldots, y_d) = \max_i |s(y_i) - s(x)|$. Clearly, Ψ is a Lipshitz function with a constant of 1.

Using the fact that F is bounded in $[0, 1]$,

$$\underset{(X,Y_1,\ldots,Y_d)\sim\mathcal{P}_{H^R}}{\mathbb{E}} \max_a |F(X) - F(Y_a)| \geqslant \underset{(X,Y_1,\ldots,Y_d)\sim\mathcal{P}_{H^R}}{\mathbb{E}} \max_a \left|\Gamma_{1-\eta}F(X) - \Gamma_{1-\eta}F(Y_a)\right| - 2\eta \tag{5.1}$$

Furthermore, since $\Gamma_{1-\eta}F$ is also bounded in $[0, 1]$, we have $s(\Gamma_{1-\eta}F) = \Gamma_{1-\eta}F$. Therefore,

$$\underset{(X,Y_1,\ldots,Y_d)\sim\mathcal{P}_{H^R}}{\mathbb{E}} \max_a \left|\Gamma_{1-\eta}F(X) - \Gamma_{1-\eta}F(Y_a)\right| = \underset{(X,Y_1,\ldots,Y_d)\sim\mathcal{P}_{H^R}}{\mathbb{E}} \max_a \left|s\left(\Gamma_{1-\eta}F(X)\right) - s\left(\Gamma_{1-\eta}F(Y_a)\right)\right| \tag{5.2}$$

Apply the invariance principle to the polynomial $Q = \left(\Gamma_{1-\eta}F, \Gamma_{1-\eta}F, \ldots, \Gamma_{1-\eta}F\right)$ and Lipshitz function Ψ. By invariance principle Theorem 5.2, we get

$$\underset{(X,Y_1,\ldots,Y_d)\sim\mathcal{P}_{H^R}}{\mathbb{E}} \max_a \left|s\left(\Gamma_{1-\eta}F(X)\right) - s\left(\Gamma_{1-\eta}F(Y_a)\right)\right|$$
$$\geqslant \underset{(Z_X,Z_{Y_1},\ldots,Z_{Y_d})\sim\mathcal{P}_{\mathcal{G}_{\Lambda,\Sigma}}}{\mathbb{E}} \max_a \left|s\left(\Gamma_{1-\eta}F(Z_X)\right) - s\left(\Gamma_{1-\eta}F(Z_{Y_a})\right)\right| - \tau^{\Omega(\varepsilon\eta/\log(1/\alpha))} \tag{5.3}$$

16

Observe that $s \circ (\Gamma_{1-\eta}F)$ is bounded in $[0, 1]$ even over the gaussian space. Hence, by using the isoperimetric result on gaussian graphs (Corollary 4.4), we know that

$$\mathop{\mathbb{E}}_{(Z_X, Z_{Y_1}, \ldots, Z_{Y_d}) \sim \mathcal{P}_{\mathcal{G}_{\Lambda, \Sigma}}} \max_a \left| s\left(\Gamma_{1-\eta}F(Z_X)\right) - s\left(\Gamma_{1-\eta}F(Z_{Y_a})\right)\right| \geqslant c \sqrt{\varepsilon \log d} \mathop{\mathbb{E}}_{Z_X, Z_Y \sim \mu_{\mathcal{G}_{\Lambda, \Sigma}}} \left| s\left(\Gamma_{1-\eta}F(Z_X)\right) - s\left(\Gamma_{1-\eta}F(Z_Y)\right)\right| \quad (5.4)$$

Now we apply the invariance principle on the polynomial $(\Gamma_{1-\eta}F, \Gamma_{1-\eta}F)$ and the functional $\Psi : \mathbb{R}^2 \to \mathbb{R}$ given by $\Psi(a, b) = |s(a) - s(b)|$. This yields,

$$\mathop{\mathbb{E}}_{Z_X, Z_Y \sim \mu_{\mathcal{G}_{\Lambda, \Sigma}}} \left| s\left(\Gamma_{1-\eta}F(Z_X)\right) - s\left(\Gamma_{1-\eta}F(Z_Y)\right)\right| \geqslant \mathbb{E}_{X, Y \sim \mu(H^R)} \left| s\left(\Gamma_{1-\eta}F(X)\right) - s\left(\Gamma_{1-\eta}F(Y)\right)\right| - \tau^{\Omega(\varepsilon\eta/\log(1/\alpha))} \quad (5.5)$$

Over H^R, the function $\Gamma_{1-\eta}F$ is bounded in $[0, 1]$, which implies that $s(\Gamma_{1-\eta}F(X)) = \Gamma_{1-\eta}F(X)$ and $\Gamma_{1-\eta}F(X) \geqslant F(X) - \eta$.

$$\mathop{\mathbb{E}}_{X, Y \sim \mu(H^R)} \left| s\left(\Gamma_{1-\eta}F(X)\right) - s\left(\Gamma_{1-\eta}F(Y)\right)\right| \geqslant \mathbb{E}_{X, Y \sim \mu(H^R)} |F(X) - F(Y)| - 2\eta \quad (5.6)$$

From equations (5.1) to (5.6) we get,

$$\mathop{\mathbb{E}}_{(X, Y_1, \ldots, Y_d) \sim \mathcal{P}_{H^R}} \max_a |F(X) - F(Y_a)| \geqslant \Omega(\sqrt{\varepsilon \log d}) \, \mathbb{E}_{X, Y \sim \mu(H^R)} |F(X) - F(Y)| - 4\eta - \tau^{\Omega(\varepsilon\eta/\log(1/\alpha))}$$

$\square$

6   Hardness Reduction from SSE

In this section we will present a reduction from SMALL-SET EXPANSION problem to BALANCED ANALYTIC VERTEX EXPANSION problem.

Let $G = (V, E)$ be an instance of SMALL-SET EXPANSION (γ, δ, M). Starting with the instance $G = (V, E)$ of SMALL-SET EXPANSION(γ, δ, M), our reduction produces an instance $(\mathcal{V}', \mathcal{P}')$ of BALANCED ANALYTIC VERTEX EXPANSION.

To describe our reduction, let us fix some notation. For a set A, let $A^{\{R\}}$ denote the set of all multisets with R elements from A. Let $G_\eta = (1 - \eta)G + \eta K_V$ where K_V denotes the complete graph on the set of vertices V. For an integer R, define $G_\eta^{\otimes R}$ to be the product graph $G_\eta^{\otimes R} = (G_\eta)^R$.

Define a Markov chain H on $V_H = \{s, t, t', s'\}$ as follows, $p(s|s) = p(s'|s') = 1 - \frac{\varepsilon}{1-2\varepsilon}$, $p(t|s) = p(t'|s') = \frac{\varepsilon}{1-2\varepsilon}$, $p(s|t) = p(s'|t') = \frac{1}{2}$ and $p(t'|t) = p(t|t') = \frac{1}{2}$. It is easy to see that the stationary distribution of the Markov chain H over V_H is given by,

$$\mu_H(s) = \mu_H(s') = \frac{1}{2} - \varepsilon \qquad \mu_H(t) = \mu_H(t') = \varepsilon$$

The reduction consists of two steps. First, we construct an "unfolded" instance $(\mathcal{V}, \mathcal{P})$ of the BALANCED ANALYTIC VERTEX EXPANSION, then we merge vertices of $(\mathcal{V}, \mathcal{P})$ to create the final output instance $(\mathcal{V}', \mathcal{P}')$. The details of the reduction are presented below.

Observe that each of the queries $\Pi(B, x)$ and $\{\Pi(C_i, y_i)\}_{i=1}^d$ are distributed according to μ' on $\mathcal{V}'$. Let $F' : \mathcal{V}' \to \{0, 1\}$ denote the indicator function of a subset for the instance. Let us suppose that

$$\mathop{\mathbb{E}}_{X, Y \sim \mathcal{V}} [|F'(X) - F'(Y)|] \geqslant \frac{1}{10}$$

Theorem 6.1. *(Completeness) Suppose there exists a set $S \subset V$ such that $\mathsf{vol}(S) = \delta$ and $\Phi(S) \leqslant \gamma$ then there exists $F' : \mathcal{V}' \to \{0, 1\}$ such that,*

$$\mathop{\mathbb{E}}_{X, Y \sim \mathcal{V}'} [|F'(X) - F'(Y)|] \geqslant \frac{1}{10}$$

and,

$$\mathop{\mathbb{E}}_{X, Y_1, \ldots, Y_d \sim \mathcal{P}} \left[\max_i |F'(X) - F'(Y_i)| \right] \leqslant 2\varepsilon + O(\eta + \gamma)$$

Proof. Define $F : \mathcal{V} \to \{0, 1\}$ as follows:

$$F(A, x) = \begin{cases} 1 & \text{if } |\Pi(A, x) \cap (S \times \{s, t\})| = 1 \\ 0 & \text{otherwise} \end{cases}$$

Observe that by definition of F, the value of $F(A, x)$ only depends on $\Pi(A, x)$. So the function F naturally

defines a map $F' : \mathcal{V}' \to \{0, 1\}$. Therefore we can write,

$$\mathbb{P}[F(A, x) = 1] = \sum_{i \in [R]} \mathbb{P}[x_i \in \{s, t\}] \, \mathbb{P}[\{a_1, \ldots, a_R\} \cap S = \{a_i\} | x_i \in \{s, t\}]$$

$$\geqslant R \cdot \frac{1}{2} \cdot \frac{1}{R} \cdot \left(1 - \frac{1}{R}\right)^{R-1} \geqslant \frac{1}{10}$$

and,

$$\mathbb{P}[F(A, x) = 1] = \mathbb{P}[|\Pi(A, x) \cap (S \times \{s, t\})| = 1] \leqslant \mathop{\mathbb{E}}_{(A,x) \sim \mathcal{V}}[|\Pi(A, x) \cap (S \times \{s, t\})|] = R \cdot \frac{1}{2} \cdot \frac{|S|}{|V|} \leqslant \frac{1}{2}$$

The above bounds on $\mathbb{P}[F(A, x) = 1]$ along with the fact that F takes values only in $\{0, 1\}$, we get that

$$\mathop{\mathbb{E}}_{X,Y \sim \mathcal{V}'}\left|F'(X) - F'(Y)\right| = \mathop{\mathbb{E}}_{(A,x),(B,y) \sim \mathcal{V}}|F(A, x) - F(B, y)| \geqslant \frac{1}{10}$$

Suppose we sample $A \in V^R$ and $B, C_1, \ldots, C_d$ independently from $G_\eta^{\otimes R}(A)$. Let us denote $A = (a_1, \ldots, a_R)$, $B = (b_1, \ldots, b_R)$, $C_i = (c_{i1}, \ldots, c_{iR})$ for all $i \in [d]$. Note that,

$$\mathbb{P}[\exists i \in [R] \text{ such that } |\{a_i, b_i\} \cap S| = 1] \quad \leqslant \sum_{i \in [R]} (1 - \eta) \, \mathbb{P}\left[(a_i, b_i) \in E[S, \bar{S}]\right] + \eta \, \mathbb{P}\left[(a_i, b_i) \in S \times \bar{S}\right]$$

$$\leqslant R(\text{vol}(S)\Phi(S) + 2\eta \, \text{vol}(S)) \leqslant 2(\gamma + \eta).$$

Similarly, for each $j \in [d]$,

$$\mathbb{P}\left[\exists i \in [R] | |\{a_i, c_{ji}\} \cap S| = 1\right] \leqslant \sum_{i \in [R]} \mathbb{P}\left[(a_i, c_{ji}) \in E[S, \bar{S}]\right] \leqslant R \, \text{vol}(S)\Phi(S) \leqslant 2(\gamma + \eta).$$

By a union bound, with probability at least $1 - 2(d + 1)(\gamma + \eta)$ we have that none of the edges $\{(a_i, b_i)\}_{i \in [R]}$ and $\{(a_i, c_{ji})\}_{j \in [d], i \in [R]}$ cross the cut $(S, \bar{S})$.

Conditioned on the above event, we claim that if $(B, x) \cap (S \times \{t, t'\}) = \emptyset$ then $\max_i |F(B, x) - F(C_i, y_i)| = 0$. First, if $(B, x) \cap (S \times \{t, t'\}) = \emptyset$ then for each $b_i \in S$ the corresponding $x_i \in \{s, s'\}$. In particular, this implies that for each $b_i \in S$, either all of the pairs $(b_i, x_i), \{(c_{ji}, y_{ji})\}_{j \in [d]}$ are either in $S \times \{s, t\}$ or $S \times \{s', t'\}$, thereby ensuring that $\max_i |F(B, x) - F(C_i, y_i)| = 0$.

From the above discussion we conclude,

$$\mathop{\mathbb{E}}_{(B,x),(C_1,y_1),\ldots,(C_d,y_d) \sim \mathcal{P}}\left[\max_i |F(B, x) - F(C_i, y_i)|\right] \leqslant \mathbb{P}[|(B, x) \cap (S \times \{t, t'\})| \geqslant 1] + 2(d + 1)(\gamma + \eta)$$

$$\leqslant \mathbb{E}[|(B, x) \cap (S \times \{t, t'\})|] + 2(d + 1)(\gamma + \eta)$$

$$= R \cdot \text{vol}(S) \cdot \varepsilon + 2(d + 1)(\gamma + \eta) = \varepsilon + 2(d + 1)(\gamma + \eta)$$

$\square$

Let $F' : \mathcal{V}' \to \{0, 1\}$ be a subset of the instance $(\mathcal{V}', \mathcal{P}')$. Let us define the following notation.

$$\text{val}_{\mathcal{P}'}(F') \stackrel{\text{def}}{=} \mathop{\mathbb{E}}_{(X,Y_1,\ldots,Y_d) \sim \mathcal{P}'}\left[\max_{i \in [d]} \left|F'(X) - F'(Y_i)\right|\right] \qquad \text{Var}_1[F'] \stackrel{\text{def}}{=} \mathop{\mathbb{E}}_{X,Y \sim \mathcal{V}'}\left|F'(X) - F'(Y)\right|$$

We define the functions $F : \mathcal{V} \to [0, 1]$ and $f_A, g_A : V_H^R \to [0, 1]$ for each $A \in V^R$ as follows.

$$F(A, x) \stackrel{\text{def}}{=} F'(\Pi(A, x)) \qquad f_A(x) \stackrel{\text{def}}{=} F(A, x) \qquad g_A(x) \stackrel{\text{def}}{=} \mathop{\mathbb{E}}_{B \sim G_\eta^{\otimes R}(A)} F(B, x)$$

Lemma 6.2.
$$\mathrm{val}_{\mathcal{P}'}(F') \geqslant \mathop{\mathbb{E}}_{A \in V^R} \mathrm{val}_{\mu_H^R}(g_A)$$

Proof.

$$\mathrm{val}_{\mathcal{P}'}(F') = \mathrm{val}_{\mathcal{P}}(F)$$

$$= \mathop{\mathbb{E}}_{A \sim V^R} \mathop{\mathbb{E}}_{x \sim \mu_H^R} \mathop{\mathbb{E}}_{y_1,\ldots,y_d \sim \mu_H^R(x)} \mathop{\mathbb{E}}_{B,C_1,\ldots,C_d \sim G_\gamma^{\otimes R}(A)} \max_i |F(B,x) - F(C_i,y_i)|$$

$$\geqslant \mathop{\mathbb{E}}_{A \sim V^R} \mathop{\mathbb{E}}_{x \sim \mu_H^R} \mathop{\mathbb{E}}_{y_1,\ldots,y_d \sim \mu_H^R(x)} \max_i \left| \mathop{\mathbb{E}}_{B \sim G_\gamma^{\otimes R}(A)} F(B,x) - \mathop{\mathbb{E}}_{C_i \sim G_\gamma^{\otimes R}(A)} F(C_i,y_i) \right|$$

$$\geqslant \mathop{\mathbb{E}}_{A \sim V^R} \mathop{\mathbb{E}}_{x \sim \mu_H^R} \mathop{\mathbb{E}}_{y_1,\ldots,y_d \sim \mu_H^R(x)} \max_i |g_A(x) - g_A(y_i)|$$

$$= \mathop{\mathbb{E}}_{A \in V^R} \mathrm{val}_{\mu_H^R}(g_A)$$

$\square$

Lemma 6.3.
$$\mathop{\mathbb{E}}_{A \sim V^R} \mathop{\mathbb{E}}_{x \sim \mu_H^R} g_A(x)^2 \geqslant \mathop{\mathbb{E}}_{(A,x) \sim \mathcal{V}} F^2(A,x) - \mathrm{val}_{\mathcal{P}'}(F')$$

Proof.

$$\mathop{\mathbb{E}}_{A \sim V^R} \mathop{\mathbb{E}}_{x \sim \mu_H^R} g_A(x)^2 = \mathop{\mathbb{E}}_{A \sim V^R} \mathop{\mathbb{E}}_{x \sim \mu_H^R} \mathop{\mathbb{E}}_{B,C \sim G_\eta^{\otimes R}(A)} F(B,x)F(C,x)$$

$$= \frac{1}{2} \mathop{\mathbb{E}}_{A \sim V^R} \mathop{\mathbb{E}}_{x \sim \mu_H^R} \mathop{\mathbb{E}}_{B,C \sim G_\eta^{\otimes R}(A)} F^2(B,x) + F^2(C,x) - (F(B,x) - F(C,x))^2$$

$$= \mathop{\mathbb{E}}_{A \sim V^R} \mathop{\mathbb{E}}_{x \sim \mu_H^R} F^2(A,x) - \frac{1}{2} \mathop{\mathbb{E}}_{A \sim V^R} \mathop{\mathbb{E}}_{x \sim \mu_H^R} \mathop{\mathbb{E}}_{B,C \sim G_\eta^{\otimes R}(A)} (F(B,x) - F(C,x))^2 \qquad (6.1)$$

where in the last step we used the fact that B,C have the same distribution as $A \sim V^R$. Since the function F is bounded in $[0,1]$, we have

$$\mathop{\mathbb{E}}_{A \sim V^R} \mathop{\mathbb{E}}_{x \sim \mu_H^R} \mathop{\mathbb{E}}_{B,C \sim G_\eta^{\otimes R}(A)} (F(B,x) - F(C,x))^2 \leqslant \mathop{\mathbb{E}}_{A \sim V^R} \mathop{\mathbb{E}}_{x \sim \mu_H^R} \mathop{\mathbb{E}}_{B,C \sim G_\eta^{\otimes R}(A)} |F(B,x) - F(C,x)| \qquad (6.2)$$

$$\mathop{\mathbb{E}}_{A \sim V^R} \mathop{\mathbb{E}}_{x \sim \mu_H^R} \mathop{\mathbb{E}}_{B,C \sim G_\eta^{\otimes R}(A)} |F(B,x) - F(C,x)|$$

$$\leqslant \mathop{\mathbb{E}}_{A \sim V^R} \mathop{\mathbb{E}}_{x \sim \mu_H^R} \mathop{\mathbb{E}}_{y \sim \mu_H^R(x)} \mathop{\mathbb{E}}_{B,C,D \sim G_\eta^{\otimes R}(A)} |F(B,x) - F(D,y)| + |F(C,x) - F(D,y)|$$

$$= 2 \mathop{\mathbb{E}}_{A \sim V^R} \mathop{\mathbb{E}}_{x \sim \mu_H^R} \mathop{\mathbb{E}}_{y \sim \mu_H^R(x)} \mathop{\mathbb{E}}_{B,D \sim G_\eta^{\otimes R}(A)} |F(B,x) - F(D,y)| \quad \text{(because (B,D), (C,D) have same distribution)}$$

$$\leqslant 2 \mathop{\mathbb{E}}_{A \sim V^R} \mathop{\mathbb{E}}_{x \sim \mu_H^R} \mathop{\mathbb{E}}_{y_1,\ldots,y_d \sim \mu_H^R(x)} \mathop{\mathbb{E}}_{B,D_1,\ldots,D_d \sim G_\eta^{\otimes R}(A)} \max_i |F(B,x) - F(D_i,y_i)|$$

$$= 2\,\mathrm{val}_{\mathcal{P}}(F) = 2\,\mathrm{val}_{\mathcal{P}'}(F') \qquad (6.3)$$

Equations (6.1), (6.2) and (6.3) yield the desired result. $\square$

Lemma 6.4.

$$\mathop{\mathbb{E}}_{A \sim V^R} \mathsf{Var}_1[g_A] = \mathop{\mathbb{E}}_{A \sim V^R} \mathop{\mathbb{E}}_{x,y \in \mu_H^R} |g_A(x) - g_A(y)| \geq \frac{1}{2}(\mathsf{Var}_1[F])^2 - \mathrm{val}_{\mathcal{P}'}(F')$$

Proof. Since the function g_A is bounded in $[0, 1]$ we can write

$$\mathop{\mathbb{E}}_{A \sim V^R} \mathop{\mathbb{E}}_{x,y \in \mu_H^R} |g_A(x) - g_A(y)| \geq \mathop{\mathbb{E}}_{A \sim V^R} \mathop{\mathbb{E}}_{x,y \in \mu_H^R} (g_A(x) - g_A(y))^2$$

$$\geq \mathop{\mathbb{E}}_{A \sim V^R} \mathop{\mathbb{E}}_{x \in \mu_H^R} g_A^2(x) - \mathop{\mathbb{E}}_{A} \mathop{\mathbb{E}}_{x,y \in \mu_H^R} g_A(x)g_A(y) \qquad (6.4)$$

In the above expression there are two terms. From Lemma 6.3, we already know that

$$\mathop{\mathbb{E}}_{A \sim V^R} \mathop{\mathbb{E}}_{x \in \mu_H^R} g_A^2(x) \geq \mathop{\mathbb{E}}_{(A,x) \sim \mathcal{V}} F^2(A, x) - \mathrm{val}_{\mathcal{P}'}(F') \qquad (6.5)$$

Let us expand out the other term in the expression.

$$\mathop{\mathbb{E}}_{A} \mathop{\mathbb{E}}_{x,y \in \mu_H^R} g_A(x)g_A(y) = \mathop{\mathbb{E}}_{A} \mathop{\mathbb{E}}_{B,C \sim G_\eta^{\otimes R}(A)} \mathop{\mathbb{E}}_{x,y \in \mu_H^R} F'(\Pi(B, x))F'(\Pi(C, y)) \qquad (6.6)$$

Now consider the following graph $\mathcal{H}$ on $\mathcal{V}'$ defined by the following edge sampling procedure.

– Sample $A \in V^R$, and $x, y \in \mu_H^R$.
– Sample independently $B \sim G_\eta^{\otimes R}(A)$ and $C \sim G_\eta^{\otimes R}(A)$
– Output the edge $\Pi(B, x)$ and $\Pi(C, y)$

Let λ denote the second eigenvalue of the adjacency matrix of the graph $\mathcal{H}$.

$$\mathop{\mathbb{E}}_{A} \mathop{\mathbb{E}}_{B,C \sim G_\eta^{\otimes R}(A)} \mathop{\mathbb{E}}_{x,y \in \mu_H^R} F'(\Pi(B, x))F'(\Pi(C, y)) = \langle F', \mathcal{H}F' \rangle$$

$$\leq \left(\mathop{\mathbb{E}}_{(A,x) \sim \mathcal{V}} F'(\Pi(A, x)) \right)^2 + \lambda \left(\mathop{\mathbb{E}}_{(A,x) \sim \mathcal{V}} (F'(\Pi(A, x)))^2 - \left(\mathop{\mathbb{E}}_{(A,x) \sim \mathcal{V}} F'(\Pi(A, x)) \right)^2 \right)$$

$$= \lambda \mathop{\mathbb{E}}_{(A,x) \sim \mathcal{V}} F(A, x)^2 + (1 - \lambda)\left(\mathop{\mathbb{E}}_{(A,x) \sim \mathcal{V}} F(A, x) \right)^2 \quad \text{(because } F'(\Pi(A, x)) = F(A, x)\text{)}$$

Using the above inequality with equations (6.4), (6.5), (6.6) we can derive the following,

$$\mathop{\mathbb{E}}_{A \sim V^R} \mathop{\mathbb{E}}_{x,y \in \mu_H^R} |g_A(x) - g_A(y)| \geq \mathop{\mathbb{E}}_{A \sim V^R} \mathop{\mathbb{E}}_{x \in \mu_H^R} g_A^2(x) - \mathop{\mathbb{E}}_{A} \mathop{\mathbb{E}}_{x,y \in \mu_H^R} g_A(x)g_A(y)$$

$$\geq (1 - \lambda)\left[\mathop{\mathbb{E}}_{(A,x) \sim \mathcal{V}} F^2(A, x) - \left(\mathop{\mathbb{E}}_{(A,x) \sim \mathcal{V}} F(A, x) \right)^2 \right] - \mathrm{val}_{\mathcal{P}'}(F')$$

$$\geq (1 - \lambda)\mathsf{Var}[F] - \mathrm{val}_{\mathcal{P}'}(F')$$

$$\geq (1 - \lambda)(\mathsf{Var}_1[F])^2 - \mathrm{val}_{\mathcal{P}'}(F') \quad \text{(because } \mathsf{Var}[F] > \mathsf{Var}_1[F]^2 \text{ for all } F\text{)}$$

To finish the argument, we need to bound the second eigenvalue λ for the graph $\mathcal{H}$. Here we will present a simple argument showing that the second eigenvalue λ for the graph $\mathcal{H}$ is strictly less than $\frac{1}{2}$. Let us restate the procedure to sample edges from $\mathcal{H}$ slightly differently.

– Define a map $\mathcal{M} : V \times V_H \to (V \cup \bot) \times (V_H \cup \{\bot\})$ as follows, $\mathcal{M}(b, x) = (b, x)$ if $x \in \{s, t\}$ and $\mathcal{M}(b, x) = (\bot, \bot)$ otherwise. Let $\Pi' : ((V \cup \bot) \times (V_H \cup \bot))^R \to (V \times \{s, t\})^{\{R\}}$ denote the following map.

$$\Pi'(B', x') = \{(b'_i, x'_i) | x_i \in \{s, t\}\}$$

– Sample $A \in V^R$ and $x, y \in \mu_H^R$

– Sample independently $B = (b_1, \ldots, b_R) \sim G_\eta^{\otimes R}(A)$ and $C = (c_1, \ldots, c_R) \sim G_\eta^{\otimes R}(A)$.

– Let $\mathcal{M}(B, x), \mathcal{M}(C, y) \in ((V \cup \{\bot\}) \times (V_H \cup \{\bot\}))^R$ be obtained by applying $\mathcal{M}$ to each coordinate of (B, x) and (C, y).

– Output an edge between $(\Pi'(\mathcal{M}(B, x)), \Pi'(\mathcal{M}(C, y)))$.

It is easy to see that the above procedure also samples the edges of $\mathcal{H}$ from the same distribution as earlier. Note that Π' is a projection from $((V \cup \bot) \times (V_H \cup \bot))^R$ to $(V \times \{s, t\})^{\{R\}}$. Therefore, the second eigenvalue of the graph $\mathcal{H}$ is upper bounded by the second eigenvalue of the graph on $((V \cup \bot) \times (V_H \cup \{\bot\}))^R$ defined by $\mathcal{M}(B, x) \sim \mathcal{M}(C, y)$. Let $\mathcal{H}_1$ denote the graph defined by the edges $\mathcal{M}(B, x) \sim \mathcal{M}(C, y)$. Observe that the coordinates of $\mathcal{H}_1$ are independent, i.e., $\mathcal{H}_1 = \mathcal{H}_2^R$ for a graph $\mathcal{H}_2$ corresponding to each coordinate of $\mathcal{M}(B, x)$ and $\mathcal{M}(C, y)$. Therefore, the second eigenvalue of $\mathcal{H}_1$ is at most the second eigenvalue of $\mathcal{H}_2$. The Markov chain $\mathcal{H}_2$ on $(V \cup \{\bot\}) \times (V_H \cup \bot)$ is defined as follows,

– Sample $a \in V$ and two neighbors $b \sim G_\eta(a)$ and $c \sim G_\eta(a)$.

– Sample $x, y \in V_H$ independently from the distribution μ_H.

– Output an edge between $\mathcal{M}(b, x)$ $\mathcal{M}(c, y)$.

Notice that in the Markov chain $\mathcal{H}_2$, for every choice of $\mathcal{M}(b, x)$ in $(V \cup \{\bot\}) \times (V_H \cup \bot)$, with probability at least $\frac{1}{2}$, the other endpoint $\mathcal{M}(c, y) = (\bot, \bot)$. Therefore, the second eigenvalue of $\mathcal{H}_2$ is at most $\frac{1}{2}$, giving a bound of $\frac{1}{2}$ on the second eigen value of $\mathcal{H}$. $\qquad \square$

Now we restate a claim from [RST12] that will be useful for our our soundness proof.

Theorem 6.5. *(Restatment of Lemma 6.11 from [RST12]) Let G be a graph with a vertex set V. Let a distribution on pairs of tuples (A, B) be defined by $A \sim V^R$, $B \sim G_\eta^{\otimes R}(A)$. Let $\ell : V^R \to [R]$ be a labelling such that over the choice of random tuples and two random permutations π_A, π_B*

$$\mathop{\mathbb{P}}_{A \sim V^R, B \sim G_\eta^{\otimes R}(A)} \mathop{\mathbb{P}}_{\pi_A, \pi_B} \left\{ \pi_A^{-1}(\ell(\pi_A(A))) = \pi_B^{-1}(\ell(\pi_B(B))) \right\} \geq \zeta$$

Then there exists a set $S \subset V$ with $\mathsf{vol}(S) \in \left[\frac{\zeta}{16R}, \frac{3}{\eta R} \right]$ satisfying $\Phi(S) \leq 1 - \zeta/16$.

The following lemma asserts that if the graph G is a *NO*-instance of SMALL-SET EXPANSION (γ, δ, M) then for almost all $A \in V^R$ the functions have no influential coordinates.

Lemma 6.6. *Fix $\delta = 1/R$. Suppose for all sets $S \subset V$ with $\mathsf{vol}(S) \in (\delta/M, M\delta)$, $\Phi(S) \geq 1 - \gamma$ then for all $\tau > 0$,*

$$\mathop{\mathbb{P}}_{A \sim V^R} \left[\exists i \mid \mathsf{Inf}_i[\Gamma_{1-\eta} g_A] \geq \tau \right] \leq \frac{1000}{\tau^3 \varepsilon^2 \eta^2} \cdot \max(1/M, \gamma)$$

22

Proof. For each $A \in V^R$, let $L_A = \left\{i \in [R] \mid \mathsf{Inf}_i(\Gamma_{1-\eta} f_A) > \tau/2\right\}$ and $L'_A = \left\{i \in [R] \mid \mathsf{Inf}_i(\Gamma_{1-\eta} g_A) > \tau\right\}$. Call a vertex $A \in V^R$ to be *good* if $L'_A \neq \emptyset$. By Lemma 5.4, the sum of influences of $\Gamma_{1-\eta} g_A$ is at most $\frac{1}{\varepsilon\eta} \mathsf{Var}[g_A] \leq \frac{1}{\varepsilon\eta}$. Therefore, the cardinality of L'_A is upper bounded by $|L'_A| \leq \frac{2}{\tau\varepsilon\eta}$. Similarly, the cardinality of L_A is upper bounded by $|L_A| \leq \frac{1}{\tau\varepsilon\eta}$.

The lemma asserts that at most a $\frac{1000}{\tau^3\eta^2\varepsilon^2} \cdot \max(1/M, \gamma)$ fraction of vertices are *good*. For the sake of contradiction, assume that $\mathbb{P}_{A\in V^R}\left[L'_A \neq \emptyset\right] \geq 1000 \max(1/M, \gamma)/\tau^2\varepsilon^2\eta^2$.

Define a labelling $\ell : V^R \to [R]$ as follows: for each $A \in V^R$, with probability $\frac{1}{2}$ choose a random coordinate in L_A and with probability $1/2$, choose a random coordinate in L'_A. If the sets L_A, L'_A are empty, then we choose a uniformly random coordinate in $[R]$.

Observe that for each $A \in V^R$, the function g_A is the average over bounded functions $f_B : V_H^R \to [0, 1]$, where $B \sim G_\eta^R(A)$. Fix a vertex $A \in V^R$ such that $L'_A \neq \emptyset$ and a coordinate $i \in L'_A$. In particular, we have that $\mathsf{Inf}_i[\Gamma_{1-\eta} g_A] \geq \tau$. Using convexity of influences, this implies that,

$$E_{B\sim G_\eta^{\otimes R}(A)} \mathsf{Inf}_i[\Gamma_{1-\eta} f_B] \geq \tau.$$

Specifically, this implies that for at least a $\frac{\tau}{2}$ fraction of the neighbours $B \sim G_\eta^R(A)$, the influence of the ith coordinate on f_B is at least $\frac{\tau}{2}$. Hence, if $L'_A \neq \emptyset$ then for at least a $\tau/2$ fraction of neighbours $B \sim G_\eta^{\otimes R}(A)$ we have $L'_A \cap L_B \neq \emptyset$.

By definition of the functions f_A, g_A, it is clear that for every permutation $\pi : [R] \to [R]$, $f_A(\pi(x)) = f_{\pi(A)}(x)$ and $g_A(\pi(x)) = g_{\pi(A)}(x)$. Therefore, for every permutation $\pi : [R] \to [R]$ and $A \in V^R$,

$$L_A = \pi^{-1}(L_{\pi(A)}) \qquad \text{and} \qquad L'_A = \pi^{-1}(L'_{\pi(A)})$$

From the above discussion, for every *good* vertex $A \in V^R$, for at least a $\tau/2$ fraction of the vertices $B \sim G_\eta^{\otimes R}(A)$, and every pair of permutations $\pi_A, \pi_B : [R] \to [R]$, we have $\pi_A^{-1}(L'_{\pi_A(A)}) \cap \pi_B^{-1}(L_{\pi_B(B)}) \neq \emptyset$. This implies that,

$$\mathbb{P}_{A\sim V^R, B\sim G_\eta^{\otimes R}(A)} \mathbb{P}_{\pi_A, \pi_B} \left\{\pi_A^{-1}(\ell(\pi_A(A))) = \pi_B^{-1}(\ell(\pi_B(B)))\right\}$$

$$\geq \mathbb{P}_{A\sim V^R}[L'_A \neq \emptyset] \cdot \mathbb{P}_{B\sim G_\eta^{\otimes R}(A)}[L'_A \cap L_B \neq \emptyset | L'_A \neq \emptyset] \cdot \mathbb{P}\left[\pi_A^{-1}(\ell(\pi_A(A))) = \pi_B^{-1}(\ell(\pi_B(B))) \mid L'_A \cap L_B \neq \emptyset\right]$$

$$\geq \mathbb{P}_{A\sim V^R}[L'_A \neq \emptyset] \cdot \left(\frac{\tau}{2}\right) \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{|L'_A|} \frac{1}{|L_B|}$$

$$\geq \mathbb{P}_{A\sim V^R}[L'_A \neq \emptyset] \cdot \left(\frac{\tau}{2}\right) \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \left(\frac{\tau\eta\varepsilon}{2}\right)^2$$

$$\geq 16 \max(1/M, \gamma)$$

By Theorem 6.5, this implies that there exists a set $S \subset V$ with $\mathsf{vol}(S) \in [\frac{1}{MR}, \frac{3}{\eta R}]$ satisfying $\Phi(S) \leq 1 - \gamma$. A contradiction. $\qquad \square$

Finally, we are ready to show the soundness of the reduction.

Theorem 6.7. *(Soundness) For all ε, d there exists choice of M and γ, η such that the following holds. Suppose for all sets $S \subset V$ with $\mathsf{vol}(S) \in (\delta/M, M\delta)$, $\Phi(S) \geq 1 - \eta$, then for all $F' : \mathcal{V}' \to [0, 1]$ such that $\mathsf{Var}_1[F'] \geq \frac{1}{10}$, we have $\mathsf{val}_{\mathcal{P}'}(F') \geq \Omega(\sqrt{\varepsilon \log d})$*

23

Proof. Fix $\eta = \varepsilon$. We will choose τ to small enough so that the error term in the soundness of dictatorship test (Proposition 5.5) is smaller than ε. Since the least probability of any vertex in Markov chain H is ε, setting $\tau = \varepsilon^{1/\varepsilon^3}$ would suffice.

First, we know that if G is a *NO*-instance of SMALL-SET EXPANSION (γ, δ, M) then for almost all $A \in V^R$, the function g_A has no influential coordinates. Formally, by Lemma 6.6, we will have

$$\mathbb{P}_{A \sim V^R} \left[\exists i \mid \mathsf{Inf}_i[\Gamma_{1-\eta} g_A] \geqslant \tau \right] \leqslant \frac{1000}{\tau^3 \eta^2} \cdot \max(1/M, \gamma).$$

For an appropriate choice of M, γ, the above inequality implies that for all but an ε-fraction of vertices $A \in V^R$, the function g_A will have no influential coordinates.

Without loss of generality, we may assume that $\mathrm{val}_{\mathcal{P}'}(F') \leqslant \sqrt{\varepsilon \log d}$, else we would be done. Applying Lemma 6.4, we get that $\mathbb{E}_{A \in V^R} \mathsf{Var}_1[g_A] \geqslant (\mathsf{Var}_1[F])^2 - \mathrm{val}_{\mathcal{P}'}(F') \geqslant \frac{1}{200}$. This implies that for at least a $\frac{1}{400}$ fraction of $A \in V^R$, $\mathsf{Var}_1[g_A] \geqslant 1/400$. Hence for at least an $1/400 - \varepsilon$ fraction of vertices $A \in V^R$ we have,

$$\mathsf{Var}_1[g_A] \geqslant \frac{1}{400} \qquad \text{and} \qquad \max_i \mathsf{Inf}_i(\Gamma_{1-\eta}(g_A)) \leqslant \tau$$

By appealing to the soundness of the gadget (Proposition 5.5), for every such vertex $A \in V^R$, $\mathrm{val}_{\mu_H^R}(g_A) \geqslant \Omega(\sqrt{\varepsilon \log d}) - O(\varepsilon) = \Omega(\sqrt{\varepsilon \log d})$. Finally, by applying Lemma 6.2, we get the desired conclusion.

$$\mathrm{val}_{\mathcal{P}'}(F') \geqslant \mathbb{E}_{A \in V^R} \mathrm{val}_{\mu_H^R}(g_A) \geqslant \Omega(\sqrt{\varepsilon \log d})$$

$\square$

7 Reduction from *Analytic d-Vertex Expansion* to Vertex Expansion

Theorem 7.1. *A* c-*vs*-s *hardness for d-*BALANCED ANALYTIC VERTEX EXPANSION *implies a* 4 c-*vs*-s/16 *hardness for balanced symmetric-vertex expansion on graphs of degree at most D, where $D = \max\{100d/s, 2\log(1/c)\}$.*

At a high level, the proof of Theorem 7.1 has two steps.

1. We show that a c-vs-s hardness for BALANCED ANALYTIC VERTEX EXPANSION. implies a 2 c-vs-s/4 hardness for instances of BALANCED ANALYTIC VERTEX EXPANSION having uniform distribution (Proposition 7.2).

2. We show that a c-vs-s hardness for instances of *d*-BALANCED ANALYTIC VERTEX EXPANSION having uniform stationary distribution implies a 2 c-vs-s/2 hardness for balanced symmetric-vertex expansion on $\Theta(D)$-regular graphs. (Proposition 7.5).

Proposition 7.2. *A* c-*vs*-s *hardness for* BALANCED ANALYTIC VERTEX EXPANSION. *implies a* 2 c-*vs*-s/4 *hardness for instances of* BALANCED ANALYTIC VERTEX EXPANSION *having uniform distribution.*

Proof. Let $(V, \mathcal{P})$ be an instance of BALANCED ANALYTIC VERTEX EXPANSION. We construct an instance $(V', \mathcal{P}')$ as follows. Let $T = 2n^2$. We first delete all vertices i from V which have $\mu(i) < 1/2n^2$, i.e. $V \leftarrow V \setminus \{i \in V : \mu(i) < 1/2n^2\}$. Note that after this operation, the total weight of the remaining vertices is still at least $1 - 1/2n$ and the BALANCED ANALYTIC VERTEX EXPANSION can increase or decrease by at most a factor of 2. Next for each i, we introduce introduce $\lceil \mu(i)T \rceil$ copies of vertex i. We will call these vertices the cloud for vertex i and index them as (i, a) for $a \in [\mu(i)T]$.

We set the probability mass of each $(d + 1)$-tuple $((i, a), (j_1, b_1) \ldots, (j_d, b_d))$ as follows :

$$\mathcal{P}'((i, a), (j_1, b_1) \ldots, (j_d, b_d)) = \frac{\mathcal{P}(i, j_1, \ldots, j_d)}{(\mu(i)T) \cdot \Pi_{\ell=1}^d (\mu(j_\ell)T)}$$

It is easy to see that $\mu'(i, a) = 1/T$ for all vertices $(i, a) \in V'$. The analytic d-vertex expansion under a function F is given by,

$$\frac{\mathbb{E}_{((i,a),(j_1,b_1)\ldots,(j_d,b_d))\sim\mathcal{P}'} \max_\ell |F(i, a) - F(j_\ell, b_\ell)|}{\mathbb{E}_{(i,a),(j,b)\sim\mu'} |F(i, a) - F(j, b)|}$$

where $X = (i, a)$ and $Y_\ell = (j, b)$ which are sampled as follows:

1. Sample a $(d + 1)$-tuple $(i, j_1, \ldots, j_d)$ from $\mathcal{P}$.

2. Sample a uniformly at random from $1, \ldots, \mu(i)T$.

3. Sample b_ℓ uniformly at random from $\{1, \ldots, \mu(j_\ell)T\}$ for each $\ell \in [d]$.

Completeness. Suppose, $\Phi(V, \mathcal{P}) \leqslant c$. Let f be the corresponding cut function. The function $f : V \mapsto \{0, 1\}$ can be trivially extended to a function $F : V' \mapsto \{0, 1\}$ thereby certifying that $\Phi(V', \mathcal{P}') \leqslant 2c$.

Soundness. Suppose $\Phi(V, \mathcal{P}) \geqslant s$. Let $F : V' \mapsto \{0, 1\}$ be any balanced function. By convexity of absolute value function we get

$$\mathbb{E}_{((i,a),(j_1,b_1),\ldots,(j_d,b_d))\sim\mathcal{P}'} \max_\ell \left|F(i, a) - F(j_{\ell,b_\ell})\right| \geqslant \mathbb{E}_{(i,j_1,\ldots,j_d)\sim\mathcal{P}} \max_\ell \left|\mathbb{E}_a F(i, a) - \mathbb{E}_\ell F(j_\ell, b_\ell)\right|.$$

So if we define $f(i) = E_a F(i, a)$, the numerator for analytic d-vertex expansion in $(V, \mathcal{P})$ for f is only lower than the corresponding numerator for F in $(V', \mathcal{P}')$. We need to lower bound the denominator, $\mathbb{E}_{i,j\sim\mu} |f(i) - f(j)|$. The requisite lower bound follows from the following two lemmas.

Lemma 7.3.

$$\mathbb{E}_{i,j\sim\mu} |f(i) - f(j)| \geqslant \mathbb{E}_{(i,a),(j,b)\sim\mu'} |F(i, a) - F(j, b)| - \mathbb{E}_{(i,a),(i,b)\sim\mu'} |F(i, a) - F(i, b)|$$

Proof. The Lemma follows directly from the following two inequalities.

$$\mathbb{E}_{(i,a),(j,b)} |F(i, a) - F(j, b)| \leqslant \mathbb{E}_{(i,a)} |F(i, a) - f(i)| + \mathbb{E}_{(j,b)} |F(j, b) - f(j)| + \mathbb{E}_{i,j} |f(i) - f(j)| \qquad \text{(Triangle Inequality)}$$

and

$$\mathbb{E}_{i,a} |F(i, a) - f(i)| \leqslant \mathbb{E}_{i,a,b} |F(i, a) - F(i, b)|$$

□

Lemma 7.4.

$$\mathbb{E}_{i,a,b} |F(i, a) - F(i, b)| \leqslant 2 \, \text{val}_{\mathcal{P}'}(F) = 2 \mathbb{E}_{(i,a),(j_1,c_1),\ldots(j_d,c_d)\sim\mathcal{P}'} \max_\ell |F(i, a) - F(j_\ell, c_\ell)|$$

Proof. Sample $(i, j_1, \ldots, j_d) \sim \mathcal{P}$. For any neighbour (j, c) of $(i, a), (i, b)$, using the Triangle Inequality we have

$$|F(i, a) - F(i, b)| \leqslant |F(i, a) - F(j, c)| + |F(j, c) - F(i, b)|$$

Therefore,

$$
\begin{aligned}
|F(i, a) - F(i, b)| \quad &\leqslant \quad \frac{\sum_\ell |F(i, a) - F(j_\ell, c_\ell)| + \sum_\ell |F(i, b) - F(j_\ell, c_\ell)|}{d} \\
&\leqslant \quad \max_\ell |F(i, a) - F(j_\ell, c_\ell)| + \max_\ell |F(i, b) - F(j_\ell, c_\ell)|
\end{aligned}
$$

Taking expectations over the uniformly random choice of a and b from the cloud of i,

$$\mathop{\mathbb{E}}_{(i,a),(i,b)} |F(i, a) - F(i, b)| \leqslant 2 \mathop{\mathbb{E}}_{((i,a),(j_1,b_1),\ldots,(j_d,b_d)) \sim \mathcal{P}'} \max_\ell |F(i, a) - F(j_\ell, c_\ell)|$$

$\square$

Lemma 7.3 and Lemma 7.4 together show that

$$\mathop{\mathbb{E}}_{i,j} |f(i) - f(j)| \geqslant \frac{\mathbb{E}_{(i,a),(j,b)} |F(i, a) - F(j, b)|}{2}.$$

as long as the value $\mathsf{val}_{\mathcal{P}'}(F) < \mathsf{Var}_1[F]/4$. Therefore, for any $F : V' \mapsto \{0, 1\}$,

$$\frac{\mathbb{E}_{((i,a),(j_1,b_1)\ldots,(j_d,b_d)) \sim \mathcal{P}'} \max_\ell |F(i, a) - F(j_\ell, b_\ell)|}{\mathbb{E}_{(i,a),(j,b) \sim \mu'} |F(i, a) - F(j, b)|} \geqslant \frac{s}{4}.$$

Theorem 4.3 shows that the minimum value of BALANCED ANALYTIC VERTEX EXPANSION is obtained by boolean functions. Therefore, $\Phi(V', \mathcal{P}') \geqslant s/4$. $\square$

Proposition 7.5. *A* c-*vs*-s *hardness for instances of d*-BALANCED ANALYTIC VERTEX EXPANSION *having uniform stationary distribution implies a* 2 c-*vs*-s/4 *hardness for balanced symmetric-vertex expansion on* $\Theta(D)$-*regular graphs. Here* $D \geqslant \max\{100d/s, 2\log(1/c)\}$.

Proof. Let $(V', \mathcal{P}')$ be an instance of d-BALANCED ANALYTIC VERTEX EXPANSION. We construct a graph G from $(V', \mathcal{P}')$ as follows : For each vertex X we pick D neighbors by sampling D/d tuples from the marginal distribution of $\mathcal{P}'$ on tuples containing X in the first coordinate.

Let $\deg(i)$ denote the degree of vertex i, i.e. the number of vertices adjacent to vertex i in G. It is easy to see that $\deg(i) \geqslant D$ and $\mathbb{E}[\deg(i)] = 2D \; \forall i \in V(G)$. Let $L = \{i \in V(G) | \deg(i) > 4D\}$. Using Hoeffding's Inequality, we get a tight concentration for $\deg(i)$ around $2D$.

$$\mathbb{P}[\deg(i) > 4D] \leqslant e^{-D}.$$

Therefore, $\mathbb{E}[|L|] < n/e^D$. We delete these vertices from G, i.e. $V(G) \leftarrow V(G) \backslash L$. With constant probability, all remaining vertices will have their degrees in the range $[D/2, 4D]$. Also, the vertex expansion of every set will decrease by at most an additive $1/e^D$.

Completeness. Let $\Phi(V', \mathcal{P}') \leqslant c$ and let $F : V' \mapsto \{0, 1\}$ be the function corresponding to $\Phi(V', \mathcal{P}')$. Let the set S be the support of the function F. Clearly, the set S is balanced. Therefore, with constant probability, we have

$$\Phi^{\mathsf{V}}(G) \leqslant \Phi_G^{\mathsf{V}}(S) \leqslant \Phi(V', \mathcal{P}') + 1/e^D \leqslant 2c.$$

Soundness. Suppose $\Phi(V', \mathcal{P}') \geqslant s$. Let $F : V' \mapsto \{0, 1\}$ be any balanced function.

Since the max is larger than the average, we get

$$\underset{X}{\mathbb{E}} \max_{Y_i \in N_G(X)} |F(X) - F(Y_i)| \geqslant \frac{d}{D} \sum_{j=1}^{D/d} \underset{(X, Y_1, ..., Y_d) \sim \mathcal{P}}{\mathbb{E}} \max_i |F(X) - F(Y_i)|$$

By Hoeffding's inequality, we get

$$\mathbb{P}\left[\underset{X}{\mathbb{E}} \max_{Y_i \in N(X)} |F(X) - F(Y_i)| < s/4\right] \leqslant \mathbb{P}\left[\frac{d}{D} \sum_{j=1}^{D/d} \underset{(X, Y_1, ..., Y_d) \sim \mathcal{P}}{\mathbb{E}} \max_i |F(X) - F(Y_i)| < s/4\right]$$

$$\leqslant \exp\left(-n(sD/d)^2\right)$$

Here, the last inequality follows from Hoeffding's inequality over the index X. There are at most 2^n boolean functions on V. Therefore, using a union bound on all those functions we get,

$$\mathbb{P}\left[\Phi^{\mathsf{V}}(G) \geqslant s/4\right] \geqslant 1 - 2^n \exp\left(-n(sD/d)^2\right).$$

Since $D > d/s$, we get that with probability $1 - o(1)$, $\Phi^{\mathsf{V}}(G) \geqslant s/4$.

$\square$

Proof of Theorem 7.1. Theorem 7.1 follows directly from Proposition 7.2 and Proposition 7.5. $\square$

8 Hardness of Vertex Expansion

We are now ready to prove Theorem 1.3. We restate the Theorem below.

Theorem 8.1. *There exists absolute constants C, β such that for every $\varepsilon > 0$, it is* SSE *-hard to distinguish between the following two cases for a given graph G with maximum degree d.*

YES *: There exists a set $S \subset V$ such that*
$$\phi^{\mathsf{V}}(S) \leqslant \varepsilon$$

NO *: For all sets $S \subset V$,*
$$\phi^{\mathsf{V}}(S) \geqslant \sqrt{\varepsilon \log d}$$

Proof. From Theorem 6.1 and Theorem 6.7 we get that for an instance of BALANCED ANALYTIC VERTEX EXPANSION $(V, \mathcal{P})$, it is SSE -hard to distinguish between the following two cases cases:

YES :
$$\Phi(V, \mathcal{P}) \leqslant \varepsilon$$

NO :
$$\Phi(V, \mathcal{P}) \geqslant c_1 \sqrt{\varepsilon \log d}$$

Now from Theorem 7.1 we get that for a graph G, it is SSE -hard to distinguish between the following two cases cases:

Yes :

$$\Phi^{\text{V,bal}} \leqslant \varepsilon$$

No :

$$\Phi^{\text{V,bal}} \geqslant c_2 \sqrt{\varepsilon \log d}$$

We use a standard reduction from Balanced vertex expansion to vertex expansion. For the sake of completeness we give a proof of this reduction in Lemma B.2. Using this reduction, we get that for a graph G, it is SSE -hard to distinguish between the following two cases cases:

Yes :

$$\Phi^{\text{V}} \leqslant \varepsilon$$

No :

$$\Phi^{\text{V}} \geqslant c_2 \sqrt{\varepsilon \log d}$$

Finally, using the computational equivalence of vertex expansion and symmetric vertex expansion (Theorem A.1), we get that for a graph G, it is SSE -hard to distinguish between the following two cases cases:

Yes :

$$\phi^{\text{V}} \leqslant \varepsilon$$

No :

$$\phi^{\text{V}} \geqslant c_2 \sqrt{\varepsilon \log d}$$

This completes the proof of the theorem.

$\square$

9 An *Optimal* Algorithm for vertex expansion

In this section we give a simple polynomial time algorithm which outputs a set S whose vertex expansion is at most $O(\sqrt{\text{OPT} \log d})$. We restate Theorem 1.2.

Theorem 9.1. *There exists a polynomial time algorithm which given a graph $G = (V, E)$ having vertex degrees at most d, outputs a set $S \subset V$, such that $\phi^{\text{V}}(S) = O(\sqrt{\text{OPT} \log d})$.*

For an undirected graph G, Bobkov et al. [BHT00] define λ_∞ as follows.

$$\lambda_\infty \overset{\text{def}}{=} \min_x \frac{\sum_i \max_{j \sim i} (x_i - x_j)^2}{\sum_i x_i^2 - \frac{1}{n}(\sum_i x_i)^2}$$

They also prove the following Theorem.

Theorem 9.2 ([BHT00]). *For any unweighted, undirected graph G, we have*

$$\frac{\lambda_\infty}{2} \leqslant \phi^{\text{V}} \leqslant \sqrt{2\lambda_\infty}$$

Consider the following SDP relaxation of λ_∞.

SDP 9.3.

$$\mathsf{SDPval} \overset{\text{def}}{=} \min \sum_{i \in} \alpha_i$$

subject to:

$$\left\| v_j - v_i \right\|^2 \;\leqslant\; \alpha_i \qquad \forall i \in V \text{ and } \forall j \sim i$$

$$\sum_i \|v_i\|^2 - \frac{1}{n} \left\| \sum_i v_i \right\|^2 \;=\; 1$$

It's easy to see that this is a relaxation for λ_∞. We present a simple randomized rounding of this SDP which, with constant probability, outputs a set with vertex expansion at most $C \sqrt{\mathsf{OPT} \log d}$ for some absolute constant C.

Algorithm 9.4.

– *Input :* A graph $G = (V, E)$

– *Output :* A set S with vertex expansion at most $72 \sqrt{\mathsf{SDPval} \log d}$ (with constant probability).

1. Compute graph G' as in Theorem A.1, let $n = |V(G')|$.
2. Solve SDP 9.3 for graph G'.
3. Pick a random Gaussian vector $g \sim N(0,1)^n$.
4. For each $i \in [n]$, define $x_i \overset{\text{def}}{=} \langle v_i, g \rangle$.
5. Sort the x_i's in decreasing order $x_{i_1} \geqslant x_{i_2} \geqslant \ldots x_{i_n}$. Let S_j denote the set of the first j vertices appearing in the sorted order. Let l be the index such that

$$l = \mathsf{argmin}_{1 \leqslant j \leqslant n/2} \Phi^\vee(S_j).$$

6. Output the set corresponding to S_l in G.

We first prove a technical lemma which shows that we can a recover a a set with small vertex expansion from a *good* linear-ordering (Step 3 in Algorithm 9.4).

Lemma 9.5. *For any* $y_1, y_2, \ldots, y_n \in \mathbb{R}^+ \cup \{0\}$, *let* $Y \overset{\text{def}}{=} [y_1 y_2 \ldots y_n]^T$ *and* $\alpha \overset{\text{def}}{=} \frac{\sum_i \max_{j \sim i} |y_j - y_i|}{\sum_i y_i}$. *Then* $\exists S \subseteq \mathsf{supp}(Y)$ *such that* $\phi^\vee(S) \leqslant \alpha$. *Morover, such a set can be computed in polynomial time.*

Proof. W.l.o.g we may assume that $y_1 \geqslant y_2 \geqslant \ldots \geqslant y_n \geqslant 0$. Then

$$\frac{\sum_i \max_{j \sim i, j < i} (y_j - y_i)}{\sum_i y_i} \leqslant \alpha \tag{9.1}$$

and

$$\frac{\sum_i \max_{j \sim i, j > i} (y_i - y_j)}{\sum_i y_i} \leqslant \alpha \tag{9.2}$$

Let $i_{\max} \overset{\text{def}}{=} \text{argmax}_i y_i > 0$, i.e. $i_{\max}$ be the largest index such that $y_{i_{\max}} > 0$. Let $S_i \overset{\text{def}}{=} \{y_1, \ldots, y_i\}$. Suppose $\forall i < i_{\max} N^v(S_i) > \alpha |S_i|$.

Now, from Inequality 9.2,

$$\alpha \geqslant \frac{\sum_i \max_{j \sim i, j < i}(y_j - y_i)}{\sum_i y_i} = \frac{\sum_i \max_{j \sim i, j < i} \sum_{l=j}^{l=i-1}(y_l - y_{l+1})}{\sum_i y_i} = \frac{\sum_i (y_i - y_{i+1})|N(S_i)|}{\sum_i y_i} > \alpha \frac{\sum_i (y_i - y_{i+1})|S_i|}{\sum_i y_i} = \alpha$$

Thus we get $\alpha > \alpha$ which is a contradition. Therefore, $\exists i \leqslant i_{\max}$ such that $\phi^V(S_i) \leqslant \alpha$. $\qquad \square$

Next we show a λ_∞-like bound for the x_i's.

Lemma 9.6. *Let $x_1, \ldots, x_n$ be as defined in Algorithm 9.4. Then, with constant probability, we have*

$$\frac{\sum_i \max_{j \sim i}(x_i - x_j)^2}{\sum_i x_i^2 - \frac{1}{n}\left(\sum_i x_i\right)^2} \leqslant 12 \, \mathsf{SDPval} \log d.$$

Proof. We will make use of the following fact that is part of the folkore about Gaussian random variables. For the sake of completeness, we prove this Fact in Appendix B (Fact B.3).

Fact 9.7. *Let $Y_1, Y_2, \ldots, Y_d$ be d normal random variables with mean 0 and variance at most σ^2. Let Y be the random variable defined as $Y \overset{\text{def}}{=} \max \{Y_i | i \in [d]\}$. Then*

$$\mathbb{E}[Y] \leqslant 2\sigma \sqrt{\log d}$$

Now using this fact we get,

$$\mathbb{E}\left[\max_{j \sim i}(x_j - x_i)^2\right] = \mathbb{E}\left[\max_{j \sim i}\langle v_i - v_j, g\rangle^2\right] \leqslant 2 \max_{j \sim i} \left\|v_j - v_i\right\|^2 \log d.$$

Therefore, $\mathbb{E}\left[\sum_i \max_{j \sim i}(x_j - x_i)^2\right] \leqslant 2 \, \mathsf{SDPval} \log d$. Using Markov's Inequality we get

$$\mathbb{P}\left[\sum_i \max_{j \sim i}(x_j - x_i)^2 < 6 \, \mathsf{SDPval} \log d\right] > \frac{2}{3}$$

For the denominator, using linearity of expectation, we get

$$\mathbb{E}\left[\sum_i x_i^2 - \frac{1}{n}\left(\sum_i x_i\right)^2\right] = \sum_i \|v_i\|^2 - \frac{1}{n}\left\|\sum_i v_i\right\|^2.$$

But since each term in the denominator is a Gaussian random variable, we get that

$$\mathbb{P}\left[\sum_i x_i^2 - \frac{1}{n}\left(\sum_i x_i\right)^2 \geqslant \frac{1}{2}\right] \geqslant \frac{1}{2}.$$

Using a union bound on the two events, we get that

$$\mathbb{P}\left[\frac{\sum_i \max_{j \sim i}(x_i - x_j)^2}{\sum_i x_i^2 - \frac{1}{n}\left(\sum_i x_i\right)^2} \leqslant 12 \, \mathsf{SDPval} \log d\right] > \frac{1}{6}.$$

$\qquad \square$

We are now ready to complete the proof of Theorem 1.2.

Proof of Theorem 1.2. Let the x_i's be as defined in Algorithm 9.4. W.l.o.g, we may assume that[1] $\left|\mathsf{supp}(x^+)\right| < \left|\mathsf{supp}(x^-)\right|$. For each $i \in [n]$, we define $y_i = x_i^+$.

Lemma 9.6 shows that with constant probability we have

$$\frac{\sum_i \max_{j \sim i}(x_i - x_j)^2}{\sum_i x_i^2 - \frac{1}{n}\left(\sum_i x_i\right)^2} \leqslant 12 \, \mathsf{SDPval} \log d.$$

We need to show that

$$\frac{\sum_i \max_{j \sim i}\left|y_i^2 - y_j^2\right|}{\sum_i y_i^2 - \frac{1}{n}\left(\sum_i y_i\right)^2} \leqslant 6 \sqrt{\frac{\sum_i \max_{j \sim i}(x_i - x_j)^2}{\sum_i x_i^2 - \frac{1}{n}\left(\sum_i x_i\right)^2}}.$$

This fact is proved in [BHT00]. For the sake of completeness, we give a proof of this fact in Appendix B (Lemma B.1). Using Lemma 9.6, we get

$$\frac{\sum_i \max_{j \sim i}\left|y_i^2 - y_j^2\right|}{\sum_i y_i^2 - \frac{1}{n}\left(\sum_i y_i\right)^2} \leqslant 72 \sqrt{\mathsf{SDPval} \log d}.$$

From Lemma 9.5 we get that the set output in Step 3 of Algorithm 9.4 has vertex expansion at most $72 \sqrt{\mathsf{SDPval} \log d}$. □

References

[ABS10] Sanjeev Arora, Boaz Barak, and David Steurer, *Subexponential algorithms for unique games and related problems*, FOCS, 2010.

[AKK+08] Sanjeev Arora, Subhash Khot, Alexandra Kolla, David Steurer, Madhur Tulsiani, and Nisheeth K. Vishnoi, *Unique games on expanding constraint graphs are easy: extended abstract*, STOC (Richard E. Ladner and Cynthia Dwork, eds.), ACM, 2008, pp. 21–28.

[Alo86] Noga Alon, *Eigenvalues and expanders*, Combinatorica **6** (1986), no. 2, 83–96.

[AM85] Noga Alon and V. D. Milman, λ_1, *isoperimetric inequalities for graphs, and superconcentrators*, J. Comb. Theory, Ser. B **38** (1985), no. 1, 73–88.

[AMS07] Christoph Ambühl, Monaldo Mastrolilli, and Ola Svensson, *Inapproximability results for sparsest cut, optimal linear arrangement, and precedence constrained scheduling*, FOCS, IEEE Computer Society, 2007, pp. 329–337.

[AR98] Yonatan Aumann and Yuval Rabani, *An O(log k) approximate min-cut max-flow theorem and approximation algorithm*, SIAM J. Comput. **27** (1998), no. 1, 291–301.

[ARV04] Sanjeev Arora, Satish Rao, and Umesh V. Vazirani, *Expander flows, geometric embeddings and graph partitioning*, STOC (László Babai, ed.), ACM, 2004, pp. 222–231.

---

[1] For any $x \in \mathbb{R}$, $x^+ \overset{\text{def}}{=} \max\{x, 0\}$.

[BHT00]   Sergey Bobkov, Christian Houdré, and Prasad Tetali, λ_∞ *vertex isoperimetry and concentration*, Combinatorica **20** (2000), no. 2, 153–172.

[FHL08]   Uriel Feige, MohammadTaghi Hajiaghayi, and James R. Lee, *Improved approximation algorithms for minimum weight vertex separators*, SIAM J. Comput. **38** (2008), no. 2, 629–657.

[IM12]   Marcus Isaksson and Elchanan Mossel, *New maximally stable gaussian partitions with discrete applications*, To Appear in Israel Journal of Mathematics, 2012.

[LLR95]   Nathan Linial, Eran London, and Yuri Rabinovich, *The geometry of graphs and some of its algorithmic applications*, Combinatorica **15** (1995), no. 2, 215–245.

[LR99]   Frank Thomson Leighton and Satish Rao, *Multicommodity max-flow min-cut theorems and their use in designing approximation algorithms*, J. ACM **46** (1999), no. 6, 787–832.

[RS10]   Prasad Raghavendra and David Steurer, *Graph expansion and the unique games conjecture*, STOC (Leonard J. Schulman, ed.), ACM, 2010, pp. 755–764.

[RST12]   Prasad Raghavendra, David Steurer, and Madhur Tulsiani, *Reductions between expansion problems*, IEEE Conference on Computational Complexity, IEEE, 2012, pp. 64–73.

[RT12]   Prasad Raghavendra and Ning Tan, *Approximating csps with global cardinality constraints using sdp hierarchies*, SODA (Yuval Rabani, ed.), SIAM, 2012, pp. 373–387.

[ST12]   David Steurer and Prasad Tetali, *Personal communication*.

[Vem05]   Santosh Vempala, *Geometric random walks: a survey*, MSRI volume on Combinatorial and Computational Geometry (2005).

[VLM10]   S Vempala, Kamal Lodaya, and Meena Mahajan, *Recent progress and open problems in algorithmic convex geometry*, IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2010), vol. 8, Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2010, pp. 42–64.

A   Reduction between Vertex Expansion and Symmetric Vertex Expansion

In this section we show that the computation of the vertex expansion is essentially equivalent to the computation of symmetric vertex expansion. Formally, we prove the following theorem.

Theorem A.1. *Given a graph G, there exists a graph G′ such that $\max_{i \in V(G)} \deg(i) = \max_{i \in V(G')} \deg(i)$ and $\phi^V(G) = \Theta(\Phi^V(G'))$. Moreover, such a G′ can be computed in time polynomial in the size of G.*

Proof. Given graph G, we construct G' as follows. We start with $V(G') = V(G) \cup E(G)$, i.e., G' has a vertex for each vertex in G and for each edge in G. For each edge $\{u, v\} \in E(G)$, we add edges $\{u, \{u, v\}\}$ and $\{v, \{u, v\}\}$ in G'. For a vertex $i \in V(G) \cap V(G')$, we set its weight to be $w(i)$. For a vertex $\{u, v\} \in E(G) \cap V(G')$, we set its weight to be $\min \{w(u)/ \deg(u), w(v)/ \deg(v)\}$.

It is easy to see that G' can be computed in time polynomial in the size of G, and that $\max_{i \in V(G)} \deg(i) = \max_{i \in V(G')} \deg(i)$.

We first show that $\phi^{\mathsf{V}}(G) \geqslant \Phi^{\mathsf{V}}(G')/2$. Let $S \subset V(G)$ be the set having the least vertex expansion in G. Let

$$S' = S \cup \{\{u, v\} \mid \{u, v\} \in E(G) \text{ and } u \in S \text{ or } v \in S\} .$$

By construction, we have $w(S) \leqslant w(S')$, $N_G(S) = N_{G'}(S')$ and

$$w(N_{G'}(\bar{S}')) \leqslant \sum_{u \in N_{G'}(S')} \deg(u) \frac{w(u)}{\deg(u)} \leqslant w(N_{G'}(S')) .$$

Therefore,

$$\Phi^{\mathsf{V}}(G') \leqslant \Phi_{G'}^{\mathsf{V}}(S') = \frac{w(N_{G'}(S')) + w(N_{G'}(\bar{S}'))}{w(S')} \leqslant \frac{2w(N_G(S))}{w(S)} = 2\phi_G^{\mathsf{V}}(S) = 2\phi^{\mathsf{V}}(G) .$$

Now, let $S' \subset V(G')$ be the set having the least value of $\Phi_{G'}^{\mathsf{V}}(S')$ and let $\varepsilon = \Phi_{G'}^{\mathsf{V}}(S')$. We construct the set S as follows. We let $S_1 = S' \backslash N_{G'}(\bar{S}')$, i.e. we obtain S_1 from S' by deleting it's internal boundary. Next we set $S = S_1 \cap V(G)$. More formally, we let S be the following set.

$$S = \left\{ v \in S' \cap V(G) \mid v \notin N_{G'}(\bar{S}') \right\} .$$

By construction, we get that $N_G(S) \subseteq N_{G'}(S') \cup N_{G'}(\bar{S}')$. Now, the internal boundary of S' has weight at most $\varepsilon w(S')$. Therefore, we have

$$w(S_1) \geqslant (1 - \varepsilon)w(S') .$$

We need a lower bound on the weight of the set S we constructed. To this end, we make the following observation. For each vertex $\{u, v\} \in S_1 \cap E(G)$, u or v also has to be in S_1 (If not, then deleting $\{u, v\}$ from S' will result in a decrease in the vertex expansion thereby contradicting the optimality of the choice of the set S'). Therefore, we have the following

$$\sum_{\{u,v\} \in S_1 \cap E(G)} w(\{u, v\}) = \sum_{\{u,v\} \in S_1 \cap E(G)} \min\left\{ \frac{w(u)}{\deg(u)}, \frac{w(u)}{\deg(u)} \right\} \leqslant \sum_{u \in S_1 \cap V(G)} w(u) = w(S) .$$

Therefore,

$$w(S) \geqslant \frac{w(S_1)}{2} \geqslant (1 - \varepsilon)\frac{w(S')}{2}$$

Therefore, we have

$$\phi^{\mathsf{V}}(G) \leqslant \phi_G^{\mathsf{V}}(S) = \frac{w(N_G(S))}{w(S)} \leqslant \frac{w(N_{G'}(S') \cup N_{G'}(\bar{S}'))}{(1 - \varepsilon)w(S')/2} = 4\Phi_{G'}^{\mathsf{V}}(S') = 4\Phi^{\mathsf{V}}(G') .$$

Putting these two together, we have

$$\frac{\phi^{\mathsf{V}}(G)}{2} \leqslant \Phi^{\mathsf{V}}(G') \leqslant 4\phi^{\mathsf{V}}(G) .$$

$\square$

B Omitted Proofs

Lemma B.1 ([BHT00]). *Let $z_1, \ldots, z_n \in R$ and let $x_i \stackrel{\text{def}}{=} z_i^+$. Then*

$$\frac{\sum_i \max_{j \sim i} \left| x_i^2 - x_j^2 \right|}{\sum_i x_i^2 - \frac{1}{n} \left(\sum_i x_i \right)^2} \leqslant 6 \sqrt{\frac{\sum_i \max_{j \sim i} (z_i - z_j)^2}{\sum_i z_i^2 - \frac{1}{n} \left(\sum_i z_i \right)^2}}.$$

Proof. W.l.o.g we may assume that $\left| \mathsf{supp}(Z^+) \right| = \left| \mathsf{supp}(Z^-) \right| = \lceil n/2 \rceil$ and that $z_1 \geqslant z_2 \geqslant \ldots \geqslant z_n$.

Note that for any $i \in [n]$, we have $\max_{j \sim i, \& j < i} (z_j^+ - z_i^+)^2 + \max_{j \sim i, \& j > i} (z_j^- - z_i^-)^2 \leqslant 2 \max_{j \sim i} (z_j - z_i)^2$. Now,

$$
\begin{aligned}
\frac{\sum_i \max_{j \sim i} (z_j - z_i)^2}{\sum_i z_i^2} \quad &\geqslant \quad \frac{\sum_i \max_{j < i \& j \sim i} (z_j^+ - z_i^+)^2 + \sum_i \max_{j > i \& j \sim i} (z_j^- - z_i^-)^2}{2 \left(\sum_{i \in \mathsf{supp}(Z^+)} z_i^2 + \sum_{i \in \mathsf{supp}(Z^-)} z_i^2 \right)} \\[2ex]
&\geqslant \quad \min \left\{ \frac{\sum_i \max_{j < i \& j \sim i} (z_j^+ - z_i^+)^2}{2 \sum_{i \in \mathsf{supp}(Z^+)} z_i^2}, \frac{\sum_i \max_{j > i \& j \sim i} (z_j^- - z_i^-)^2}{2 \sum_{i \in \mathsf{supp}(Z^-)} z_i^2} \right\}
\end{aligned}
$$

W.l.o.g we may assume that

$$\frac{\sum_i \max_{j < i \& j \sim i} (z_j^+ - z_i^+)^2}{\sum_{i \in \mathsf{supp}(Z^+)} z_i^2} \leqslant \frac{\sum_i \max_{j > i \& j \sim i} (z_j^- - z_i^-)^2}{\sum_{i \in \mathsf{supp}(Z^-)} z_i^2}$$

$$\frac{\sum_i \max_{j \sim i} (x_j - x_i)^2}{\sum_i x_i^2} \leqslant 2 \frac{\sum_i \max_{j \sim i} (z_j - z_i)^2}{\sum_i z_i^2}$$

We have

$$
\begin{aligned}
\max_{j \sim i, j < i} (x_j^2 - x_i^2) \quad &= \quad \max_{j \sim i, j < i} (x_j - x_i)(x_j + x_i) \\[1ex]
&\leqslant \quad \max_{j \sim i, j < i} \left((x_j - x_i)^2 + 2 x_i (x_j - x_i) \right) \\[1ex]
&\leqslant \quad \max_{j \sim i, j < i} (x_j - x_i)^2 + 2 x_i \max_{j \sim i, j < i} (x_j - x_i) \\[1ex]
&\leqslant \quad \sum_i \max_{j \sim i, j < i} (x_j - x_i)^2 + 2 \sqrt{\sum_i x_i^2} \sqrt{\max_{j \sim i, j < i} (x_j - x_i)^2} \qquad \text{Cauchy-Schwarz} \\[1ex]
&= \quad \lambda_\infty \sum_i x_i^2 + 2 \sqrt{\lambda_\infty} \sum_i x_i^2
\end{aligned}
$$

Thus we have

$$\frac{\sum_i \max_{j \sim i, j < i} (x_j^2 - x_i^2)}{\sum_i x_i^2} \leqslant 6 \sqrt{\frac{\sum_i \max_{j \sim i} (z_j - z_i)^2}{\sum_i z_i^2}}$$

$\square$

Lemma B.2. *A* c*-vs-*s *hardness for b-Balanced-vertex expansion implies a* 2 c*-vs-*s/2 *hardness for vertex expansion.*

Proof. Fix a graph $G = (V, E)$.

Completeness. If G has Balanced-vertex expansion at most c, then clearly its vertex expansion is also at most c.

Soundness. Suppose we have a polynomial time algorithm that outputs a set S having $\phi^{\vee}(S) \leqslant s$ whenever G has a set S' having $\phi^{\vee}(S') \leqslant 2c$. Then this algorithm can be used as an oracle to find a balanced set of vertex expansion at most s. This would contradict the hardness of Balanced-vertex expansion.

First we find a set, say T, having $\phi^{\vee}(T) \leqslant s$. If we are unable to find such a T, we stop. If we find such a set T and T has balance at least b, then we stop. Else, we delete the vertices in T from G and repeat. We continue until the number of deleted vertices first exceeds a $b/2$ fraction of the vertices.

If the process deletes less than $b/2$ fraction of the vertices, then the remaining graph (which has at least $(1 - b/2)n$ vertices) has conductance $2c$, and thus in the original graph every b-balanced cut has conductance at least c. This is a contradiction !

If the process deletes between $b/2$ and $1/2$ of the nodes, then the union of the deleted sets gives a set T' with $\phi^{\vee}(T') \leqslant s$ and balance of T' at least $b/2$.

Therefore, starting with $b = 2b$, we arrive at a contradiction. $\qquad\square$

Fact B.3. *Let $Y_1, Y_2, \ldots, Y_d$ be d standard normal random variables. Let Y be the random variable defined as $Y \stackrel{\text{def}}{=} \max\{Y_i | i \in [d]\}$. Then*

$$\mathbb{E}[Y] \leqslant 2\sqrt{\log d}$$

Proof. For any $Z_1, \ldots, Z_d \in \mathbb{R}$ and any $p \in \mathbb{Z}^+$, we have $\max_i |Z_i| \leqslant (\sum_i Z_i^p)^{\frac{1}{p}}$. Now $Y^2 = (\max_i X_i)^2 \leqslant \max_i X_i^2$.

$$
\begin{aligned}
\mathbb{E}[Y^2] \;\; &\leqslant\;\; \mathbb{E}\left[\left(\sum_i X_i^{2p}\right)^{\frac{1}{p}}\right] \leqslant \left(\mathbb{E}\left[\sum_i X_i^{2p}\right]\right)^{\frac{1}{p}} \quad (\text{ Jensen's Inequality }) \\
&\leqslant\;\; \left(\sum_i \left(\mathbb{E}[X_i^2]\right)\frac{(2p)!}{(p)!2^p}\right)^{\frac{1}{p}} \leqslant 2pd^{\frac{1}{p}} \quad (\text{using } (2p)!/p! \leqslant (2p)^p)
\end{aligned}
$$

Picking $p = \log d$ gives $\mathbb{E}[Y^2] \leqslant 2e \log d$.
Therefore $\mathbb{E}[Y] \leqslant \sqrt{\mathbb{E}[Y^2]} \leqslant \sqrt{2e \log d}$.

$\qquad\square$

C Noise Operators

Let H be a Markov chain and let $F : V(H^k) \mapsto \{0, 1\}$ be any boolean function. In this section we prove some basic properties of $\Gamma_{1-\eta}F$. We restate the definition of our Noise Operator $\Gamma_{1-\eta}$.

$$\Gamma_{1-\eta}F(X) = (1 - \eta)F(X) + \eta \mathop{\mathbb{E}}_{Y \sim X} F(Y)$$

The Fourier expansion of the function F is $F = \sum_\sigma \hat{f}_\sigma e_\sigma$ where $\{e_\sigma\}$ is the set of eigenvectors of H^k. It is easy to see that $e_\sigma = e_{\sigma_1} \otimes \ldots \otimes e_{\sigma_k}$, where the $\{e_{\sigma_i}\}$ are the eigenvectors of H.

Lemma C.1. *(Decay of High degree Coefficients) Let Q_j be the multi-linear polynomial representation of* $\left|\Gamma_{1-\eta}F(X) - \Gamma_{1-\eta}F(Y_j)\right|$. *Then,*

$$\mathsf{Var}(Q_j^{>p}) \leqslant (1 - \varepsilon\eta)^{2p}$$

Proof.

$$
\begin{aligned}
\Gamma_{1-\eta}F(X) &= (1-\eta)F(X) + \eta \mathop{\mathbb{E}}_{Y\sim X} F(Y) \\
&= \sum_\sigma \hat{f}_\sigma \, \mathbb{E}\left[e_\sigma(X) + \mathop{\mathbb{E}}_{Y\sim X} F(Y)\right] \\
&= \sum_\sigma \hat{f}_\sigma \Pi_{i\in\sigma}\left((1-\eta)e_{\sigma_i}(X_i) + \mathop{\mathbb{E}}_{Y_i\sim X_i} e_{\sigma_i}(Y_i)\right)
\end{aligned}
$$

We bound the second moment of $\Gamma_{1-\eta}F$ as follows

$$
\begin{aligned}
\mathop{\mathbb{E}}_X\left(\Gamma_{1-\eta}F(X)\right)^2 &= \sum_\sigma \hat{f}_\sigma^2 \mathop{\mathbb{E}}_X \Pi_{i\in\sigma}\left((1-\eta)e_{\sigma_i}(X_i) + \eta \mathop{\mathbb{E}}_{Y_i\sim X_i} e_{\sigma_i}(Y_i)\right)^2 \\
&= \sum_\sigma \hat{f}_\sigma^2 \Pi_{i\in\sigma}\left[(1-\eta)^2 \mathop{\mathbb{E}}_{X_i} e_{\sigma_i}(X_i)^2 + \eta^2 \mathop{\mathbb{E}}_{X_i}\left(\mathop{\mathbb{E}}_{Y_i\sim X_i} e_{\sigma_i}(Y_i)\right)^2 + 2\eta(1-\eta)\mathop{\mathbb{E}}_{X_i}\mathop{\mathbb{E}}_{Y_i\sim X_i} e_{\sigma_i}(X_i)e_{\sigma_i}(Y_i)\right)^2 \\
&= \sum_\sigma \hat{f}_\sigma^2 \Pi_{i\in\sigma}\left((1-\eta)^2 + \eta^2\lambda_i^2 + 2\eta(1-\eta)\lambda_i\right) \\
&= \sum_\sigma \hat{f}_\sigma^2 \Pi_{i\in\sigma}(1-\eta+\eta\lambda_i)^2
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\mathsf{Var}(Q_j^{>p}) &\leqslant 4\sum_{\sigma:|\sigma|>p} \hat{f}_\sigma^2 \Pi_{i\in\sigma}(1-\eta+\eta\lambda_i)^2 \\
&\leqslant \sum_{\sigma:|\sigma|>p} \hat{f}_\sigma^2 (1-\varepsilon\eta)^{2|\sigma|} \\
&\leqslant (1-\varepsilon\eta)^{2p}
\end{aligned}
$$

Here the second inequality follows from the fact that all non-trivial eigenvalues of H are at most $1 - \varepsilon$ and the third inequality follows Parseval's indentity. $\qquad\square$